

UNIVERSITÉ D'ALGER
FACULTÉ MIXTE DE MÉDECINE ET DE PHARMACIE D'ALGER
ANNÉE 1938 — N° 8

Contribution à l'Etude
des
STATISTIQUES MÉDICALES
LE ROLE DU HASARD
LA NOTION D'ÉCART PROBABLE

THÈSE

pour le Doctorat en Médecine
présentée et soutenue publiquement le 3 Juin 1938
par

Emile RICARD

NÉ LE 13 JUIN 1909 A SAINT-EUGÈNE (DÉP. D'ALGER)
ANCIEN INTERNE DES HOPITAUX
ANCIEN AIDE D'ANATOMIE A LA FACULTÉ DE MÉDECINE D'ALGER
LAURÉAT DE LA FACULTÉ

JURY

MM. SENEVET, <i>Professeur de parasitologie et Zoologie médicale</i>	Président
PINOY, <i>Professeur de bactériologie</i>	Juges
PORTES, <i>Professeur de physique médicale</i>	
BENHAMOU, <i>Agrégé (Médecine)</i>	

UNIVERSITÉ D'ALGER
FACULTÉ MIXTE DE MÉDECINE ET DE PHARMACIE D'ALGER
ANNÉE 1938 — N° 8

Contribution à l'Etude
des
STATISTIQUES MÉDICALES
LE ROLE DU HASARD
LA NOTION D'ÉCART PROBABLE

THÈSE

pour le Doctorat en Médecine
présentée et soutenue publiquement le 3 Juin 1938
par

Emile RICARD

NÉ LE 13 JUIN 1909 A SAINT-EUGÈNE (DÉP. D'ALGER)
ANCIEN INTERNE DES HOPITAUX
ANCIEN AIDE D'ANATOMIE A LA FACULTÉ DE MÉDECINE D'ALGER
LAURÉAT DE LA FACULTÉ

JURY

MM. SENEVET, <i>Professeur de parasitologie et Zoo-</i> <i>logie médicale</i>	Président
PINOY, <i>Professeur de bactériologie</i>	} Juges
PORTES, <i>Professeur de physique médicale</i>	
BENHAMOU, <i>Agrégé (Médecine)</i>	

UNIVERSITÉ D'ALGER

FACULTÉ MIXTE DE MÉDECINE ET DE PHARMACIE D'ALGER

DOYEN	MM. LEBLANC (☞. ☞. I. ☞)
DOYEN HONORAIRE	HERAIL (☞. I. ☞)
ASSESEUR	COSTANTINI (* ☞. I. ☞)

PROFESSEURS

Clinique médicale	AUBRY ☞. ☞. I. ☞)
Clinique gynécologique	CABANES (☞. ☞. I. ☞)
Clinique chirurgicale	COSTANTINI (* ☞. I. ☞)
Botanique et matière médicale	FOURMENT (I. ☞. O.O. ☞)
Clinique médicale et hygiène infantile	GILLOT (I. ☞)
Médecine légale	GIRAUD (O. ☞. ☞. I. ☞)
Clinique obstétricale et puériculture du premier âge	LAFFONT (O. ☞. ☞. I. ☞)
Anatomie	LEBLANC (☞. ☞. I. ☞)
Thérapeutique	LEBON (☞. ☞. I. ☞)
Clinique chirurgicale infantile et d'orthopédie	LOMBARD (* ☞. I. ☞)
Chimie biologique et médicale	GIBERTON (I. ☞)
Pharmacie	MUSSO (* I. ☞)
Bactériologie	PINOY (* I. ☞)
Clinique psychiatrique	POROT (I. ☞)
Physique médicale	PORTES (☞. ☞. I. ☞)
Anatomie pathologique	MONTPELLIER I. ☞)
Clinique des maladies des pays chauds, des maladies syphilitiques et cutanées	RAYNAUD (☞. ☞. I. ☞)
Parasitologie et zoologie médicale	SENEVET (☞. I. ☞)
Physiologie	TOURNADE (O. ☞. ☞. I. ☞)
Histologie et Embryologie	COURRIER (* ☞. I. ☞)
Chimie générale, pharmaceutique et toxicologique	WUNSCHENDORFF (I. ☞)
Clinique ophtalmologique	TOULANT (* ☞. ☞)
Hygiène, hydrologie et climatologie	N...

PROFESSEURS HONORAIRES

ARGAUD (☞. ☞. I. ☞)	HERAIL (☞. I. ☞)
POUJOL (☞. I. ☞)	WEBER (* ☞. I. ☞)

CHARGES DE COURS

Clinique d'oto-rhino-laryngologie	MM. ABOULKER (☞. ☞. I. ☞)
---	---------------------------

AGREGES

Médecine	BENHAMOU (* ☞. A. ☞)
Histologie	LACROIX (A. ☞)
Chirurgie	N...
Obstétrique	SABADINI
Anatomie pathologique	MOGGI (* I. ☞)
Anatomie	HOUEL (* ☞. I. ☞)
Chimie médicale	N...
Histoire naturelle pharmaceutique	RIBET (☞. A. ☞)
Pharmacie	N...
	ROQUES (I. ☞)
	MONNET

AGREGES LIBRES

Physiologie	BATTEZ
Chirurgie	FERRARI (I. ☞)
	DUBOUCHER (I. ☞)

NOTA. — La Faculté n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions doivent être considérées comme propres à leurs auteurs.

SERMENT D'HIPPOCRATE

En présence des Maîtres de cette Faculté, de mes chers condisciples et selon la tradition d'Hippocrate, je promets et je jure d'être fidèle aux lois de l'honneur et de la probité dans l'exercice de la Médecine. Je donnerai mes soins gratuits à l'indigent, et n'exigerai jamais un salaire au-dessus de mon travail. Admis dans l'intérieur des maisons, mes yeux ne verront pas ce qui s'y passe, ma langue taira les secrets qui me seront confiés et mon état ne servira pas à corrompre les mœurs ni à favoriser le crime.

Respectueux et reconnaissant envers mes Maîtres, je rendrai à leurs enfants l'Instruction que j'ai reçue de leurs pères.

Que les hommes m'accordent leur estime si je suis fidèle à mes promesses. Que je sois couvert d'opprobre et méprisé de mes confrères si j'y manque.



Digitized by the Internet Archive
in 2026

Je dédie cette thèse :

A CEUX DES MIENS QUI NE SONT PLUS

A MA FEMME. CHARLETTE RICARD

Docteur en Médecine

*Hommage de mon immense
amour.*

A MA GRAND'MERE

A MA MERE ET A MON PERE

*Témoignage de ma profonde
affection.*

A MON FRERE, LE DOCTEUR GEORGES RICARD

A MES TANTES ET A MES ONCLES

A MES BEAUX-PARENTS

A MA BELLE-SŒUR ANNE-MARIE RICARD

A MA BELLE-SŒUR

LE DOCTEUR MARCELINE ANTOINE

A MON PRESIDENT DE THESE

MONSIEUR LE PROFESSEUR SENEVET

*Qui a inspiré ce travail et
nous fait l'honneur de prési-
der aujourd'hui notre jury.*

A MES JUGES

MONSIEUR LE PROFESSEUR PINOY

MONSIEUR LE PROFESSEUR PORTES

MONSIEUR LE PROFESSEUR AGRÉGÉ BENHAMOU

Qui furent aussi nos Maîtres

A MON MAITRE LE DOCTEUR RIBET

Professeur agrégé d'Anatomie

Nous n'oublierons jamais la bienveillance avec laquelle il nous a reçu au Laboratoire d'Anatomie.

Nous garderons toujours le souvenir de son enseignement si clair.

Nous aurions aimé pouvoir poursuivre l'étude de l'anatomie avec un maître tel que lui.

Des circonstances fâcheuses ne l'ont pas permis...

Qu'il veuille bien trouver ici la respectueuse assurance de toute notre reconnaissance et de notre entier dévouement.

A MONSIEUR LE PROFESSEUR AGRÉGÉ BENHAMOU

Nous le remercions de la sympathie qu'il n'a cessé de nous témoigner au cours de notre internat.

Ses admirables leçons resteront toujours gravées dans notre mémoire.

Qu'il veuille bien trouver ici, un faible témoignage de notre gratitude et de notre profond respect.

A MADEMOISELLE LE DOCTEUR RENÉE ANTOINE

Ancien Chef de Clinique à la Faculté

A MONSIEUR LE DOCTEUR DERRIEU

Médecin des Hôpitaux

MONSIEUR LE DOCTEUR PIERRE GOINARD

Chirurgien des Hôpitaux

A MES AMIS

MONSIEUR LE DOCTEUR MORALI .

MONSIEUR LE DOCTEUR CLAUDE SOLACROUP

MONSIEUR LE DOCTEUR ANDRÉ MONPERE

MONSIEUR LE DOCTEUR BALLISTE

MONSIEUR LE DOCTEUR CALLEJA

MONSIEUR LE DOCTEUR SY

MONSIEUR LE DOCTEUR BECMEUR

MONSIEUR LE DOCTEUR HOMAR

MONSIEUR LE DOCTEUR LIARAS

Ancien Prosecteur

MONSIEUR LE DOCTEUR EZES

Ancien Aide d'Anatomie

MONSIEUR LE DOCTEUR SEROR

Ancien Aide d'Anatomie

MONSIEUR LE DOCTEUR LACOSTE

MONSIEUR LE MÉDECIN-CAPITAINE DELABIE

MONSIEUR CAMILLE FAUCHIER

Docteur en Pharmacie

MADAME ET MONSIEUR RENÉ BAYLE

Docteur en Pharmacie

MADemoiselle LE DOCTEUR HÉLÈNE CORIZZI

MONSIEUR LE DOCTEUR MINICONI

MONSIEUR LE DOCTEUR FLOTTES

MONSIEUR LE DOCTEUR

FRIMIGACCI-STEPHANOPOLI

MONSIEUR LE DOCTEUR AMARANTINI

MONSIEUR LE DOCTEUR FONTENEAU

Contribution à l'Étude
des
STATISTIQUES MÉDICALES
LE RÔLE DU HASARD
LA NOTION D'ÉCART PROBABLE

AVANT-PROPOS

On a beaucoup ironisé sur les statistiques : « l'art de préciser ce qu'on ignore », « l'art de faire mentir les mathématiques », « l'art d'effectuer des additions exactes avec des chiffres inexacts », « un des trois types de mensonge »... etc...

Il est vrai que les statistiques ne sont pas exemptes de critiques, mais c'est encore le seul moyen d'information par lequel nous puissions nous faire une idée des résultats obtenus grâce aux procédés mis en œuvre contre la maladie et les fléaux sociaux.

Sans statistique pas de prophylaxie possible, plus d'expériences comparatives sans le moyen de dresser le bilan des variations statistiques.

Il est donc nécessaire d'établir des statistiques, mais encore faut-il qu'elles soient établies sur des bases rationnelles et exactes. Il faut ensuite les interpréter.

Le problème qui se pose au sujet des statistiques médicales est donc double :

- 1°. — Comment doit-on bâtir une statistique ?
- 2°. — Comment doit-on l'interpréter ?

Ces deux parties sont évidemment intimement liées et elles se confondent souvent. Il est évident que l'interprétation exacte d'une statistique dépendra avant tout de la rigueur scientifique des procédés

employés pour la bâtir. Cependant on peut scinder les deux problèmes et n'envisager que l'interprétation.

Une statistique normalement établie, toute cause d'erreur étant écartée, il reste encore un facteur qui peut faire varier dans d'énormes limites son interprétation. Ce facteur c'est le hasard.

Notre travail se propose d'étudier ce « facteur hasard » dans l'interprétation des statistiques médicales, laissant volontairement à d'autres l'étude des principales causes d'erreur qui grèvent lourdement leur établissement. (1)

Nous préciserons tout d'abord ce que nous entendons par le rôle du hasard dans l'interprétation d'une statistique. Nous étudierons ensuite les lois qui le régissent, le calcul des probabilités, qui nous permet de calculer sa participation dans les événements que comporte la statistique, la façon dont on utilise ce calcul dans la recherche des marges d'erreur probables et inévitables, par les tables d'erreur et les nomogrammes à points alignés, et dans le problème des concomitances.

Enfin et pour permettre l'emploi du calcul des probabilités aux praticiens qui répugnent à l'usage des mathématiques nous analyserons la méthode des « statistiques témoins hétérogènes », d'AUGUSTE LUMIERE, ainsi qu'une méthode purement expérimentale.

(1) Un travail sur les erreurs de base des statistiques médicales est sur le point d'être terminé par Madame le Docteur Charlette Ricard.

OBSERVATION N° 1

*(Exemple d'erreur par la méconnaissance de la notion
d'écart probable.)*

Observations tirée des Archives des Instituts Pasteur
de l'Afrique du Nord.

(Tome VI 1928, p. 18 à 22)

L'INFECTION TUBERCULEUSE REVELEE PAR LES CUTI-REACTIONS A LA TUBERCULINE CHEZ LES INDIGENES DE FIGUIG

(Maroc Oriental)

Par J. MATHIEU

« Profitant des tournées de vaccination jennérienne que nous avons faite dans les sept ksour du pays de Figuig : El Hammam, Tahtani, El Hammam Fougani, El Maïz, Oudaghir, les Abidat, les Ouled Sliman et, Zenaga, nous avons recueilli 1.012 observations de cuti-réactions à la tuberculine, chez les Berbères blancs, les Haratins, et les Israélites de la région. Nous avons opéré suivant la méthode préconisée par L. Parrot et F. Foley, dans une étude parue dans les archives des Instituts Pasteur de l'Afrique du Nord (T. II, fas. I, Mars 1922, p. 72-78).

Il nous a paru intéressant de comparer les résultats statistiques que nous avons obtenus avec ceux des enquêtes faites par l'Institut Pasteur d'Algérie

de 1911 à 1914 (H. Foley) (1), et en 1924 (Brouard et Leblanc) chez les israélites de Beni Ounif et de Figuig. (2)

L'épreuve de cuti réaction à la tuberculine chez les indigènes musulmans de Figuig nous a donné pendant les mois d'Octobre et de Novembre 1927, les résultats suivants :

1. — Résultats obtenus en 1927 par J. Mathieu.

TABLEAU I

Maroc Oriental, FIGUIG, Population berbère :

Enfants de 1 à 15 ans :

138 cuti + sur 692 sujets = 19,94 %

Adultes :

30 cuti + sur 52 sujets = 57,69 %

Index tuberculinique total : = 38,81 %

TABLEAU II

Maroc Oriental, FIGUIG, Population négroïde :

Enfants de 1 à 15 ans :

5 cuti + sur 37 sujets = 13,5 %

Adultes :

10 cuti + sur 16 sujets = 62,5 %

Index tuberculinique total : = 38 %

TABLEAU III

Maroc Oriental FIGUIG, Population israélite :

-
- (1) L. PARROT et H. FOLET. — L'infection tuberculeuse chez les indigènes musulmans d'Algérie. Archives des Instituts Pasteur de l'Afrique du Nord, Tome II, Fas. I, Mars 1922 (p. 79-109).
- (2) M. BROUARD et L. LEBLANC. — L'infection tuberculeuse chez les israélites de Beni-Ounif et de Figuig. Archives de l'Institut Pasteur d'Algérie, Tome II, Fas. 4, Décembre 1924 (p. 478-484).

Enfants de 1 à 15 ans :

7 cuti + sur 78 sujets = 8,9 %

Adultes :

37 cuti + sur 65 sujets = 56,9 %

Index tuberculinique total : = 32,9 %

2. — Résultats donnés par H. Foley concernant les cuti réactions à la tuberculine faites de 1911 à 1914.

Chez les berbères blancs :

Enfants de 1 à 15 ans : 17,7 %

Adultes : 38,4 %

Index total : 28 %

Chez les négroïdes :

Enfants de 1 à 15 ans : 0 %

Adultes : 11 cuti + sur 15

Index total : 36,5 %

3. — Résultats donnés par L. Leblanc en 1923 pour la population israélite.

Chez les israélites :

Enfants de 1 à 15 ans :

6 cuti + sur 73 sujets = 8 %

Adultes :

47 cuti + sur 87 sujets = 56 %

Index total : = 30,5 %

4. — Tableau comparatif des différents pourcentages obtenus entre 1911 et 1927, à Figuig, chez les berbères blancs, les négroïdes et les israélites.

Berbères blancs

	1 à 15	Adultes	I. T.
de 1911 à 1914	17,7 %	38,4 %	28 %
en 1927	19,4 %	57,9 %	38,81 %

Négroïdes

	1 à 15	Adultes	I. T.
de 1911 à 1914	0 %	11/15	36,5 %
en 1927	13,5 %*	62,5 %	38 %

Israélites

	1 à 15	Adultes	I. T.
en 1923	8 %	56 %	30,5 %
en 1927	8,9 %	56,9 %	32,9 %

M. J. Mathieu s'empresse alors de conclure :

« Ces comparaisons semblent démontrer une augmentation dans le degré de l'infection tuberculeuse chez les indigènes de l'oasis de Figuig... »

Puis s'appuyant sur cette première affirmation qui nous le démontrerons plus loin est des plus téméraires, M. J. Mathieu cherche les causes de cette augmentation de l'infection tuberculeuse à Figuig. L'accès plus facile, les chemins de fer, les automobiles, les mines de Bou Arfa, l'alcool (« ils ne dédaignent pas le verre d'anisette bu entre amis à l'abri des regards indiscrets. »), sont tour à tour incriminés dans l'accroissement du nombre des cuti réactions positives.

Et M. J. Mathieu en arrive à dire : « l'infection tuberculeuse latente s'exprime à Figuig selon une loi générale, par une courbe ascendante d'allure comparable à la courbe qui exprimerait l'européanisation croissante d'une région. »

L'utilisation des Tables de Ross et Scott ou du nomogramme de Millot, nous permettent d'affirmer à la probabilité de 1/1000 que les chiffres de 30,5 % et de 32,9 % donnés par L. Leblanc en 1923 et par J. Ma-

thieu en 1927 comme index tuberculinique total de la population israélite de Figuig sont approchés à + ou — 12 %, les index totaux sont donc compris entre :

18,5 % et 42,5 % pour 1923

20,9 % et 44,9 % pour 1927

Ce qui signifie que lorsqu'on n'examine dans une population que 160 sujets ou 143, le pourcentage d'erreur est de + ou — 12 %, à la probabilité de 1/1000.

Nous pouvons, en utilisant les Tables seules, dire qu'il aurait fallu examiner 40.936 individus pour affirmer 1 contre 100.000, que le chiffre de 30,5 % était approché à + ou — 1 %, c'est-à-dire, compris entre 29,5 % et 31,5 %.

Les moyens que nous avons employés pour faire ces critiques seront développés aux Chapitre III.

OBSERVATION N° 2

(Exemple d'erreur par la méconnaissance du rôle du hasard dans les distributions inégales des faits.)

Observation tirée de la Presse Médicale du 19 Mai 1920, par J. Herber

LE PALUDISME AU CAMP DE.... (Maroc) SES CAUSES (1)

« Cette note n'a pour but que l'interprétation d'un document rigoureusement exact.

Chargé du service médical d'un poste pendant plus d'un an, je l'ai établie d'après des constatations faites pendant les quatre premiers mois de ma mission. Cette période fut suffisante pour étudier la répartition des cas de paludisme dans ce camp, car mon arrivée fut antérieure à l'invasion de la maladie et mon observation prit fin à sa disparition.

Dans ce poste, les tentes étaient au nombre de 34 ; elles contenaient chacune 8 à 10 hommes ; les cas de paludisme se répartirent ainsi que le figure le schéma ci-après, où chaque grand carré représente une tente, chaque petit carré un homme atteint ; la position respective des tentes a été très exactement reproduite.

(1) D'après un rapport adressé à la direction du Service de Santé en 1916 sur le paludisme au camp de..... en 1915.

Le simple examen de ce croquis montre que les tentes furent inégalement atteintes « s'il s'agissait de toute autre maladie on n'hésiterait pas à parler de contagion », mais le paludisme étant en jeu, on n'admet plus qu'un groupement forfuit.

Personnellement, je ne crois pas que le hasard puisse tout expliquer: sur les 5 tentes qui contenaient de 4 à 6 malades, 2 abritaient des hommes venant de pays très impaludés, 3 d'un pays réputé sain.

Ces tentes étaient exactement semblables, et si irrégulièrement réparties dans le camp, qu'on ne pouvait incriminer leur situation. Leur dissémination même, tout aussi bien que l'existence de tentes épargnées par le paludisme, rendait difficilement admissible l'hypothèse de la propagation par l'intermédiaire d'un insecte volant. Il n'y a d'ailleurs pas de moustiques dans la région.

Pendant toute la période qu'a duré cette observation, je n'en ai pas vu un seul et tous ceux qui ont eu l'attention attirée sur leur recherche, n'ont pas été plus heureux que moi. A la vérité j'ai capturé quelques mois plus tard un anophèle, mais l'eussé-je trouvé plus tôt qu'une double constatation s'imposerait, l'extrême rareté des moustiques et la fréquence du paludisme. La théorie de la transmission du paludisme par l'anophèle est une certitude, mais c'est une hypothèse de dire que le moustique est le seul agent de contagion. Pour toutes les maladies transmises par les insectes, on a recours aujourd'hui à des formules simples, puce=peste,, mouche=dysenterie, pou = typhus, moustique = paludisme. C'est d'ailleurs à la suite de la découverte de l'évolution de l'hématozoaire dans le corps de l'anophèle que cette tendance d'esprit s'est fait jour.

La connaissance de la pathologie générale nous conduit à des pensées moins exclusives : elle ne s'oppose point à la conception d'insectes, vecteurs passifs de l'hématozoaire, à côté de l'intermédiaire actif qu'est le moustique.

En voyant le dessin ci-dessous qui est le résultat de plusieurs mois d'observation, on ne peut s'empêcher de penser aux autres insectes suceurs. Sans ailes, leur zone de nocivité est restreinte ; ils peuvent être coupables de contagion de lit à lit, et créer seulement de petites épidémies de tente. A l'époque où j'ai recueilli ce document, il faisait une température de 40 à 50 degrés ; les hommes couchaient à demi nus sur des paillasses contigües et disposées sur le sol. Les fébricitants et les valides étaient dans une promiscuité absolue. Ils se trouvaient dans les conditions les plus favorables pour que les insectes les piquent consécutivement. Il n'y avait pas de punaises dans le camp, mais les puces abondaient, et, bien que la preuve de la transmission du paludisme par leur intermédiaire soit difficile, le microscope en main, il faut y songer au nom de la clinique. Leur présence peut expliquer à la fois la dissémination du paludisme dans une région où il n'y a pas de moustiques, et l'apparition de petits foyers épidémiques dans une région où habitent des paludéens.

Cette hypothèse clinique fait comprendre certains faits connus, difficilement explicables avec la théorie du moustique ; elle permet de concevoir pourquoi certaines maisons voisines identiques, se comportent différemment vis-à-vis du paludisme ; dans les unes tous les hôtes sont atteints, dans les autres ils sont indemnes. Elle n'est pas en désaccord avec les données de la pathologie générale. Elle est, en outre,

féconde ; elle conduit à des idées nouvelles pour la lutte contre le paludisme ; elle explique la formation de foyers secondaires en pays sains et elle donne le moyen de les détruire.

Il importe de la vérifier en recueillant de nouveaux faits. »

J. HERBER.



Sous le titre « Paludisme et calcul des probabilités » (éléments de biométrie, pages 169 à 175), le Docteur Martinet a discuté cette répartition inégale dans les 34 tentes impaludées. Il a établi, par le calcul des probabilités, que la répartition des paludéens dans les 34 tentes étaient due seulement au hasard, détruisant ainsi toute hypothèse sur la transmission du plasmodium par un insecte suceur.

CHAPITRE I

QU'EST-CE QUE LE ROLE DU HASARD?

QU'EST-CE QUE LE ROLE DU HASARD

Ainsi, il ne suffit donc pas d'établir correctement une statistique. Il faut l'interpréter, c'est à dire en dégager tout l'enseignement qu'elle comporte, en extraire ce qui doit diriger nos efforts vers le progrès car toute statistique est un exposé dépourvu d'explication, de vues théoriques, de tout système, et consistant en un simple inventaire.

Pour l'interpréter correctement, il faut tout d'abord savoir comment elle a été établie, puis chercher à connaître quelle est la marge d'erreur inévitable que ses résultats comportent. Puis quand on connaît cette marge d'erreur il faut se demander si les résultats qui plaident en faveur d'une causalité quelconque ne sont pas dus simplement à des coïncidences. Alors seulement la statistique aura toute la rigueur scientifique qui sied à une méthode sur laquelle sont basés tous les inventaires d'une nation, tous les progrès de la Science.

Quel est le facteur qui conditionne les marges d'erreur inévitables et qui détermine les coïncidences ?

C'est le hasard.

Il paraît à première vue, saugrenu de baser une étude que nous voulons précise sur un facteur aussi peu stable que le hasard, mais à première vue seulement, car le hasard n'est plus aussi capricieux que le voulaient les auteurs classiques ; il est devenu tout

un chapitre des Sciences mathématiques sous le nom de Calcul des probabilités.

Notre conception scientifique du hasard n'est pas celle d'une force intelligente, tantôt bienveillante, tantôt ennemie qui favorise ou contrarie de manière tout à fait imprévue et imprévisible nos désirs et nos intérêts. Ce n'est pas davantage une croyance à l'existence de faits « rebelles à toute loi », car devant les précisions et les vérifications toujours plus rigoureuses de la science, ce dernier reste de fétichisme a disparu et le principe du déterminisme reste incontesté.

Le hasard désigne simplement les événements dont la connaissance imparfaite certes nous permet simplement de prévoir, avec une approximation plus ou moins grande, les causes, les effets et les lois en s'aidant du calcul des probabilités.

Ainsi, au hasard capricieux, se substitue le calcul précis des probabilités.

CHAPITRE II

LES PRINCIPES DU CALCUL DES PROBABILITES

LES PRINCIPES DU CALCUL DES PROBABILITES

Avant que d'entreprendre l'étude du rôle du hasard par le calcul des probabilités, nous voulons rappeler ici, les travaux du grand mathématicien français Laplace qui, en 1812 écrivit sa théorie analytique des probabilités, ceux de J. Bertrand, de l'Académie Française, secrétaire perpétuel de l'Académie des Sciences, qui édita en 1917 un ouvrage sur le calcul des probabilités, ceux de Carvallo, Directeur de l'Ecole Polytechnique, qui présenta le calcul des probabilités et ses applications, ceux de Ross et Stott qui en 1911 publient des Tables dans lesquelles on trouve les probabilités de pourcentage d'erreur des différents calculs statistiques, ceux, en 1923, d'Henri Poincaré, Membre de l'Institut, qui précise quelques points sur la relativité des Sciences mathématiques et résout de nouveaux théorèmes de probabilité, ceux de Maurice d'Ocagne et de Stanislas Millot qui mettent au point un nomogramme pour l'application du théorème de Jacques Bernoulli, ceux de Stanislas Millot qui reprend l'étude de la probabilité à priori, et qui après avoir montré toutes ses difficultés, met au point sa « Théorie nouvelle de la probabilité des causes », en 1925.

Sous les auspices de ces maîtres, nous nous permettons une esquisse timide des quelques principes élémentaires du calcul des probabilités.

CALCUL DES PROBABILITES

On considère comme également possibles, c'est-à-dire comme ayant les mêmes chances d'arriver, ou ayant la même probabilité, tous les événements dont les causes ne présentent aucune différence.

Par exemple, si l'on a mis dans une urne, des boules d'égal diamètre, de même surface, de même poids mais qui se distinguent à la vue, qu'on les agite et qu'on en fasse tirer une, on dira qu'elles ont toutes la même probabilité de sortir.

Si les boules de l'urne sont les une blanches, au nombre de m , les autres noires au nombre de n , la probabilité qu'il sorte une boule blanche sera :

$$x = \frac{m}{m + n}$$

Si toutes les questions étaient aussi simples que celle-là, il n'y aurait pas de théorie des probabilités.

Le calcul des probabilités a pour but de ramener les questions compliquées de chance à ce degré de simplicité. La méthode consiste à ramener tous les événements à une même probabilité et à réduire ensuite le nombre des chances favorables et celui des chances défavorables.

PREMIER PRINCIPE

Le principe permettant d'effectuer ces transformations consiste en ce que « l'on peut multiplier ou diviser par un nombre les événements favorables et défavorables, supposés tous également possibles, sans que la probabilité d'un événement de l'une ou de l'autre espèce en soit aucunement altéré. »

Explication :

Si au lieu de m blanches et n noires, nous mettons dans l'urne mp boules blanches et np boules noires, mais que les mp boules blanches soient divisées en m groupes de p boules, offrant un caractère distinct, à la vue seulement, et que les np boules noires soient groupées de la même manière, il est évident qu'en fait, nous aurons dans l'urne m groupes blancs et n groupes noirs, présentant tous les mêmes chances d'être touchés par l'une des p boules qui les composent :

La probabilité de tirer une blanche sera donc la même que de toucher un des groupes blancs, c'est-à-dire :

$$\frac{m}{m + n}$$

DEUXIEME PRINCIPE

Il se rapporte aux événements composés.

La probabilité d'un événement composé ou la probabilité composée, est le produit des probabilités partielles.

A. — Si on met dans une urne trois blanches et sept noires, qu'on doive en tirer une, la remettre et refaire un nouveau tirage, il pourra sortir : deux blanches, deux noires, ou une noire et une blanche.

Quelle est la probabilité de la sortie de deux blanches ?

Comme une même boule peut sortir deux fois de suite, le nombre des accouplements est $10 \times 10 = 100$. Le nombre des accouplements favorables est 3×3 ou 9 donc la probabilité est :

$$\frac{9}{100} \text{ ou } \frac{3}{10} \times \frac{3}{10}$$

c'est-à-dire le carré de la probabilité simple pour la sortie d'une blanche.

Pour la probabilité de la sortie de deux noirs :

$$7 \times 7 = 49 \text{ d'où } \frac{49}{100} = \frac{7}{10} \times \frac{7}{10}$$

Pour la probabilité de la sortie d'une blanche et d'une noire :

$$3 \times \frac{7}{100} = \frac{21}{100} = \frac{3}{10} \times \frac{7}{10}$$

B. — Il y a m blanches et n noires dans une urne m' blanches et n' noires dans une autre urne, on voudrait tirer successivement, une boule blanche de l'une et de l'autre urne.

Le nombre des accouplements est :

$$(m + n) \times (m' + n')$$

d'un autre côté, le nombre des accouplements favorables est : mn'. La probabilité est donc :

$$\frac{m m'}{(m + n) (m' + n')} \text{ ou } \left(\frac{m}{m + n} \right) \left(\frac{m'}{m' + n'} \right)$$

c'est-à-dire toujours le produit des probabilités simples.

Parfois les événements sont dépendants les uns des autres.

Exemple : Dans une urne, il y a m petites urnes blanches et n noires. Dans chaque urne blanche il se

trouve m' boules blanches et n' noires, dans chaque urne noire m'' blanches et n'' noires.

Calcul de la probabilité pour qu'on tire :

1° une urne blanche, puis :

2° de cette urne blanche une boule blanche.

Solution : On ne changera ni la probabilité d'amener une urne ni celle d'en tirer ensuite une boule blanche si l'on multiplie dans toutes les urnes blanches les nombres de boules blanches et noires par $m'' + n''$ et dans toutes les urnes noires le nombre des boules blanches et noires par $m' + n'$ on aura donc : m urnes blanches et n urnes noires contenant : les premières :

$m' (m'' + n'')$ blanches et $n' (m'' + n'')$ noires

les autres :

$m'' (m' + n')$ blanches et $n'' (m' + n')$ noires

le nombre de boules dans toutes les urnes sera

$$(m' + n') \times (m'' + n'')$$

Ceci posé, tous les événements sont également possibles, puisque toutes les petites urnes contiennent le même nombre de boules. Il suffit de compter le nombre total des événements et le nombre des événements favorables :

Le nombre total des événements composés est :

$$(m + n) (m' + n') (m'' + n'')$$

puisque'il y a $(m + n)$ urnes contenant chacune

$$(m' + n') (m'' + n'')$$
 boules.

Le nombre des événements composés favorables est d'ailleurs : $mm' (m'' + n'')$ puisque'il y a m urnes blanches contenant chacune $m' (m'' + n'')$ boules blanches.

La probabilité cherchée est donc :

$$\frac{mm' (m'' + n'')}{(m + n) (m' + n') (m'' + n'')}$$

C'est-à-dire encore le produit des probabilités des événements simples : le tirage d'une urne blanche et tirage d'une boule blanche et cette urne blanche.

Cette démonstration s'étendrait sans peine au cas d'un événement, composé d'une série d'un nombre quelconque d'événement successifs se gouvernant les uns les autres. En effet, l'événement complet pourra être regardé comme composé du dernier et de l'ensemble de tous les autres. La probabilité de ce dernier étant déjà le produit des probabilités des événements simples précédents, la probabilité de l'événement total sera bien le produit de toutes les probabilités simples successives.

Ainsi : la probabilité d'un événement compose d'un nombre d'événements simples, réglés les uns par les autres, s'obtient en multipliant la probabilité que le premier étant arrivé, le deuxième arrivera, puis par la probabilité que les deux premiers étant arrivés, le troisième arrivera et ainsi de suite.

TROISIEME PRINCIPE

Il se rapporte à des événements différents, mais considérés tous comme favorables.

La probabilité totale est la somme des probabilités partielles.

Supposons qu'il y ait dans une urne :

m boules blanches, n boules bleues, p boules rouges, s boules vertes, t boules noires ;

La probabilité qu'il sorte une blanche sera :

$$\frac{m}{(m + n + p + s + t)}$$

La probabilité qu'il sorte une bleue sera :

$$\frac{n}{(m + n + p + s + t)}$$

La probabilité de ne pas amener une noire sera égale à la somme des probabilités d'amener une blanche ou une bleue ou une verte etc.... soit :

$$\frac{m}{(m+n+p+s+t)} + \frac{n}{(m+n+p+s+t)} + \frac{p}{(m+n+p+s+t)} \times \frac{s}{(m+n+p+s+t)}$$

QUATRIEME PRINCIPE

Il se rapporte à la comparaison des causes

Supposons que des causes inégalement probables, mais dont les probabilités seront données, puissent produire des nombres différents et connus d'événements, les uns favorables, les autres défavorables, mais dont un seul doit arriver, et cherchons à déterminer la probabilité d'un événement favorable.

Soient m, m', m'' les probabilités des différentes causes, p et q les nombres d'événements favorables et défavorables que la première cause peut produire, p' et q' ; p'' et q'' ; etc...., les nombres analogues relatifs à la seconde cause, à la troisième, etc...

Il est clair d'abord que la probabilité de l'événement favorable ne sera pas changée si, à la place des

causes dont les probabilités sont m, m', m'' , etc.... nous supposons m causes de la première espèce m' causes de la seconde, m'' causes de la troisième etc..., pourvu que toutes les causes soient maintenant également probables.

D'un autre côté, on ne changera pas la probabilité de l'événement favorable, quelle que soit la cause qui doit se trouver en action, si on multiplie les nombres d'événements favorables et défavorables que chacune peut produire par le produit des nombres d'événements que toutes celles d'espèces différentes peuvent produire séparément.

Or, cela fait, on n'aura plus à considérer que des causes également probables, pouvant toutes produire le même nombre d'événements ; tous les événements seront donc devenus également possibles, et par conséquent il ne restera plus qu'à compter le nombre total des événements et le nombre des événements favorables.

Soit P le produit $(q + q') \times (q'' + q''') \dots$ qui représente maintenant le nombre d'événements que peut produire une quelconque des causes, le nombre total des événements sera :

$$(m + m' + m'' + \dots) \times (P) ;$$

d'un autre côté, le nombre des événements favorables sera :

$$\frac{m \ p \ P}{p \ q} + \frac{m' \ p' \ P}{p' + q'} + \frac{m'' \ p'' \ P}{p'' + q''} + \dots ;$$

la probabilité d'un événement favorable sera donc :

$$\frac{m \ p}{(p+q) (m+m'+m''+\dots)} + \frac{m' \ p'}{(p'+q') (m+m'+m''+\dots)} + \dots ;$$

la probabilité sera donc exprimée par la somme des

produits de la probabilité de chaque cause par la probabilité qu'elle fait acquérir à l'événement favorable.

Tels sont les principes du calcul des probabilités.

Comment utilise-t-on ce calcul ?

Les auteurs se divisent en deux catégories :

- les uns préfèrent les Tables,
- les autres les Nomogrammes.

Les Tables donnent les résultats des calculs de probabilité déjà effectués par leurs auteurs.

Chacune d'elle correspond à un degré de probabilité donné, en général elles envisagent des probabilités allant de une chance sur deux ou 0,50 ou pile ou face, jusqu'à une probabilité de une chance sur 100.000 ou 0,99.999 de la nomenclature française.

Elles évaluent les pourcentages d'erreur depuis 1 % jusqu'à 6 % mais permettent par des calculs très simples de trouver les pourcentages en deçà de 1 % et au delà de 6 %.

Toutes sont construites pour servir à des pourcentages trouvés de 1 % à 100 %.

Les Nomogrammes sont des graphiques établis sur les variations de théorème de probabilités. Celui que nous reproduisons est construit pour l'application du théorème de Jacques Bernouilli.

Grâce à un fil tendu entre les différentes valeurs de pourcentages trouvés, de marges d'erreur consentie, de probabilité donnée, on trouvera, sans aucun calcul et très rapidement, l'un des points qui nous intéresse : nombre de choses à examiner, pourcentage d'erreur ou probabilité pour ce nombre ou ce pourcentage.

CHAPITRE III

LA NOTION D'ECART PROBABLE OU MARGE D'ERREUR INEVITABLE

LA NOTION D'ECART PROBABLE OU MARGE D'ERREUR INEVITABLE

Supposons que nous désirions étudier des éléments de nature quelconque, des cas de maladie, des leucocytes, des trypanosomes etc... et que nous voulions nous assurer qu'un lot des éléments considérés se distingue de l'ensemble par quelque particularité commune. On ne peut trancher la question avec une certitude absolue que si on examine l'ensemble des sujets dont a été extrait le lot considéré.

Mais c'est en général impossible, et nous devons nous contenter de l'examen d'un nombre assez restreint de ces éléments. En notant le rapport qui existe entre les éléments à caractéristiques spéciales et le nombre des sujets examinés, nous pouvons prévoir le rapport probable qui existera entre ces éléments à caractéristiques spéciales et l'ensemble de tous les éléments considérés.

Exemple : Si nous désirons savoir avec certitude combien d'individus d'une certaine nationalité ont les yeux bleus, nous devons examiner tous les individus de cette nationalité. Mais comme c'est impossible, nous n'examinerons que dix ou cent ou mille au plus de ces individus ; nous noterons combien d'entr'eux ont les yeux bleus, puis nous déduirons de cet échantillon quelle est la proportion des sujets aux yeux bleus de toute la nation.

Evidemment nous nous approcherons d'autant plus de la vérité que nous examinerons davantage de sujets; mais nous n'aurons atteint une absolue certitude qu'après avoir examiné tous les habitants du pays.

La question qui se pose est donc de savoir combien d'éléments nous devons examiner pour être sûrs que notre résultat contient un pourcentage donné de l'exacte vérité, c'est-à-dire de connaître la marge d'erreur probable que notre résultat contient.

Les travaux des mathématiciens nous permettent de résoudre ce problème et les tables d'erreur probable ainsi que les nomogrammes nous donnent la possibilité de répondre sans aucun calcul à la plupart des cas.

(A)

LE CALCUL DE L'ECART PROBABLE
PAR L'APPLICATION
DES TABLES DE ROSS ET STOTT

I. — LES TABLES

Les Tables d'erreurs probables les plus employées sont celles qui ont été dressées par Ross et Stott sous le nom de « Tables of statistical error ».

Nous reproduisons, traduite le plus près possible du texte anglais, la pensée de ces auteurs, concernant l'emploi des tables.

Le style en a beaucoup souffert mais le fond a été conservé.

Ils nous disent :

« Nous commencerons à nous occuper des cas dans lesquels le nombre total des éléments est si grand que nous ne pouvons étudier qu'une très petite proportion de ce nombre total.

I. — LES ECHANTILLONS

PROPORTIONNELLEMENT PETITS.

Tout d'abord précisons ce que nous entendons par les phrases « pourcentage donné de certitude » et « raisonnablement sûr ».

Ces deux phrases contiennent des idées de degré. Ainsi, si nous menons une enquête strictement scientifique, nous pouvons désirer que le pari (ou probabilité) soit de 99.999 chances contre une que notre résultat contienne 1 % de certitude.

Nous devons alors consulter la Table A1 sous la deuxième colonne. S'il est suffisant d'avoir une probabilité de 99.999 contre 1 que notre résultat contienne 5 % de la certitude nous consultons la même Table

A1 mais sous la sixième colonne. Il peut arriver encore qu'un plus petit degré de probabilité nous suffise, 99 contre 1 par exemple que notre résultat contienne 3 % de la vérité. Dans ce cas nous consultons la Table A4 sous la quatrième colonne. Enfin nous pouvons nous contenter de pile ou face, pur hasard, ce qui équivaut à une probabilité de 1 contre 1 et pour cela nous consultons la Table A6.

Par ces exemples, le lecteur aura compris facilement l'utilisation des Tables

1° Les chiffres qui sont placés en tête de chacune indiquent le degré de probabilité de certitude avec laquelle nous pouvons nous fier aux conclusion tirées de cette table.

Ainsi le rapport est de 99.999 pour 1 que les conclusions tirées de la première table sont exactes, ou en d'autres mots : en dehors de 100.000 résultats de cette Table un seul peut être seulement erroné vraisemblablement.

La Table A5 est vraisemblablement faussée une fois sur 10 résultats, et la Table A6 une fois sur deux résultats.

2. Les pourcentages inscrits au début de chaque colonne donnent le pourcentage de « l'erreur statistique » c'est-à-dire la quantité par laquelle la vérité absolue est vraisemblablement divergente du résultat que nous avons obtenu.

Si notre résultat est 70 % et que l'erreur statistique est de 5 % cela signifie que la vérité absolue se trouvera comprise entre 75 % et 65 %. Si l'erreur statistique est de 1 % la vérité sera entre 71 % et 69 %.

3. Si les résultats observés « plus l'erreur » excè-

dent 100, ou si le résultat « moins l'erreur » est plus petit que zéro, nous concluons que le nombre des choses que nous avons examinées n'est pas suffisamment élevé pour procurer un résultat utile.

4. Les chiffres de la première colonne, à gauche de chaque table indiquent les pourcentages observés, jusqu'à 50 %.

5. Ceux qui sont dans le corps de la table indiquent le nombre total des choses qui doivent être examinées pour nous permettre d'affirmer que le résultat obtenu est grevé du pourcentage d'erreur inscrit en tête de la colonne, avec le degré de probabilité placé au début de chaque Table.

Exemple :

Si le résultat obtenu est 43 %, nous devons examiner 47.780 choses pour affirmer que dans la probabilité de 99.999 pour un, la vérité est comprise entre 44 % et 42 %, le pourcentage d'erreur étant de 1 %.

Mais nous n'aurons à examiner que 1328 choses pour affirmer que, pour une probabilité de 99.999 pour un, la vérité se trouve comprise entre 49 % et 37 %, le pourcentage d'erreur étant de 6 %.

Nous devons examiner 26.541 choses pour affirmer, avec une probabilité de 999 pour un que la vérité est comprise entre 44 et 42 % (Table A3).

Mais nous ne devons examiner que 415 choses pour être sûr avec une probabilité de 9 pour un que la vérité est comprise entre 47 % et 39 % (Table A5).

Et seulement 185, pour que, avec une probabilité de 9 pour un, la vérité soit comprise entre 49 % et 37 %.

6. Si le pourcentage observé est plus grand que 50 % nous le soustrayons de 100 % et obtenons par le

reste de la soustraction les chiffres qui vont nous être utiles pour opérer les mêmes recherches que précédemment.

Exemple :

Le pourcentage trouvé est de 70 % ou 79 % ou 85 %. Soustraits de 100 % ces pourcentages deviennent : 30 %, 21 %, 15 %.

7. En examinant les Tables nous voyons à la fois : que les chiffres diminuent rapidement dans les tables successives et dans les colonnes successives d'une même Table.

De ces observations nous déduisons que le nombre de choses qu'il est nécessaire d'examiner diminue.

1°. — avec une probabilité de correction réduite.

2°. — avec une erreur statistique plus grande.

Les chiffres augmentent tandis que nous descendons les colonnes, cela signifie que le nombre de choses que nous devons examiner pendant que le pourcentage trouvé s'élève de 1 à 50 %, et diminue tandis que le pourcentage observé continue à croître de 50 à 100 %. D'où la règle qui permet de calculer l'erreur probable de 80 % avec 20 %.

8. Les Tables sont établies pour six degrés d'erreur statistique de 1 à 6 %.

Il est facile d'obtenir les chiffres pour quelque degré que ce soit, en divisant les chiffres de la colonne à 1 % par le carré du degré d'erreur recherché.

Ainsi pour une probabilité de 99.999 pour 1, et un pourcentage observé de 43 % nous devons examiner 478 choses pour une erreur statistique de 10 %, et 4.778.000 choses pour une erreur statistique de 0,10 %.

Les chiffres des colonnes à 2 %, 3 %, 4 %, etc... sont soumis à cette même règle.

9. il est aussi facile d'obtenir approximativement les chiffres pour des pourcentages avec décimales tels que 2,5 % et 27,3 %, car l'accroissement du nombre des choses qu'il nous faut examiner est, en gros, proportionnel à l'accroissement du pourcentage trouvé

Soustrayons le chiffre immédiatement au dessous du chiffre immédiatement au dessus, et multiplions le reste de cette soustraction par le pourcentage à décimales que nous avons choisi, additionnons le résultat de cette multiplication au chiffre le plus bas situé qui nous a servi déjà à effectuer la soustraction. Nous aurons ainsi le nombre de choses qu'il nous faudra examiner.

Ainsi dans la Table A1, dans la colonne 1 % le chiffre correspondant à 2,5 % de pourcentage à atteindre sera environ de 4.747, et celui qui correspond à 27,3 % sera environ 38.686 ; mais de telles précisions sont très rarement recherchées.

10. Nous sommes souvent obligés d'examiner un si petit nombre de choses que l'erreur statistique est évidemment plus grande que le 6 % prévu par les tab'les. Si nous voulons savoir de combien l'erreur de notre pourcentage trouvé dépasse ce chiffre il faut procéder de la **manière suivante** :

Choisissez le degré de probabilité qui convient et dans la Table appropriée pointez le pourcentage que vous avez obtenu. Prenez ensuite le nombre correspondant à ce pourcentage dans la colonne à 1 % d'erreur et diviser ce nombre par le nombre de choses que vous avez examinées.

L'erreur sera la racine carrée du quotient de cette division.

Exemple :

Imaginons que nous n'ayions examiné que 300 leucocytes dans une formule leucocytaire, et que 45 % de ces leucocytes soient des monocléaires.

Le nombre de leucocytes que nous devons examiner pour que, la probabilité étant de 9.999 pour 1, l'erreur soit de 1 % est de : 37.459. Le rapport de ce nombre à 300 est 125 et la racine carrée de $125 = 11,2$. D'où l'erreur statistique de notre travail, à une probabilité de 9.999 pour 1, est de 11,2 %.

Mais 6.697 leucocytes auraient suffi pour une probabilité de 9 pour 1, Le rapport de 6.697 à 300 est 22,32 et la racine carrée de ce nombre est de 4,7, ce qui représente l'erreur statistique cherchée pour une probabilité de 9 pour 1.

En résumé le lecteur se rendra compte du grand nombre de choses que l'on doit examiner avant de pouvoir obtenir un résultat avec un haut degré de probabilité et dans d'étroites limites d'erreur. Il se souviendra sans aucun doute de beaucoup d'assertions hardies basées sur l'examen d'un très petit nombre d'échantillons.

Exemples :

1) — Combien de gens d'une nation faut-il examiner avant de pouvoir être sûr, pour une probabilité de 99.999 pour 1 que 65 à 67 % de toute la nation n'a pas d'yeux bleus.

Réponse : 43.744.

2) — Combien de leucocytes devons nous examiner chez un malade avant de pouvoir affirmer 999 contre

un que 40 à 42 % de tous ces leucocytes sont des mononucléaires ?

Réponse : 26.194.

3) — Combien de leucocytes doivent être examinés avant de pouvoir dire 99 contre un que 65 à 69 % de tous ces leucocytes sont des polynucléaires ?

Réponse : 3.668

4) — Combien devons nous examiner de ces leucocytes pour pouvoir affirmer, pour une probabilité de 9.999 pour un, que le nombre des polynucléaires éosinophiles de 1,5 % à 2,5 % du nombre total des leucocytes ?

Réponse : 11.868 (multipliez le chiffre de la colonne à 1 % d'erreur par 4.)

5) — En examinant 100 de ces leucocytes, nous trouvons que 7 % d'entr'eux sont de grands mononucléaires. Combien de leucocytes nous devons examiner en plus pour pouvoir affirmer, 10.000 contre un, que ce même pourcentage de 7 % se maintient à tous les leucocytes du corps ?

Réponse : 995.

6) — En examinant 75 observations de cas morbides, nous trouvons que 13 des malades moururent. Combien d'observations de cas de cette maladie devons nous examiner en plus pour dire, avec une probabilité de 99 pour 1, que le pourcentage de mortalité de la maladie est compris entre 19 et 21 % ?

Réponse : 10.551.

7) — Notre nouvelle méthode de traitement a sauvé un malade sur 5 dans une maladie réputée jusqu'alors incurable. Combien devons nous traiter de cas avant d'affirmer que, pour une probabilité de 999 pour 1, nous pouvons guérir entre 15 et 25 % de tous les cas traités ?

Réponse : 688.

8) — En examinant près de 2.000 hématies chez un malade, nous trouvons que 8 % d'entr'elles sont nucléées. Quel sera le pourcentage des hématies nucléées dans le corps entier de l'individu, avec une probabilité de 9 contre 1 ?

Réponse : 7 et 9 %.

9) — En examinant 220 choses, nous trouvons que 31 d'entr'elles appartiennent à une catégorie particulière. Quelle est l'erreur statistique à la probabilité de 99 pour 1 ?

Réponse : 6 %.

10) — En examinant 138 leucocytes, nous trouvons que 15 % sont de grands mononucléaires. Quelle est la proportion des grands mononucléaires dans le corps entier, avec une probabilité d'environ 1000 à 1 ?

Réponse : un pourcentage compris entre 5 et 25 %.

11) — En examinant 500 parasites de malaria, nous trouvons que 14 % ont des formes sexuées. Quelle est la probabilité que l'erreur statistique soit d'environ 1 % ?

Réponse : une chance égale.

12) — Dans le même cas, quelle est la probabilité que l'erreur statistique ne soit pas plus grande que 4 %.

Réponse : 99 pour 1.

13) — Sur 200 leucocytes, nous trouvons 23 % de mononucléaires. Quelle est l'erreur statistique pour une probabilité de 999 pour 1 ?

Réponse : 9,79 % (racine carrée du quotient de 19.178 par 200.)

14) — Le jour suivant chez le même patient, sur le

même nombre de leucocytes, nous trouvons 40 % de mononucléaires. Quelle est l'erreur statistique pour la même probabilité ?

Réponse : 11,4 %.

15) — Pouvons nous parier 999 contre 1, qu'il y a un accroissement de mononucléaires dans ce cas ?

Réponse : Non, car les erreurs se chevauchent : la différence entre 23 et 40 est plus petite que la somme $9,79 + 11,4$ %.

16) — Pouvons nous alors parier qu'il y a eu un accroissement, avec une probabilité de 99 pour 1 ?

Réponse : Oui, parce qu'avec un degré de probabilité plus bas, la somme des erreurs, à savoir: $7,7 + 8,9$ % est plus petite que la différence entre les pourcentages observés 23 et 40 %.

17) — Faisant un travail avec une probabilité de 999 pour 1, nous trouvons que sur 100 leucocytes, 2 sont éosinophiles, soit 2 %. Quel est le pourcentage d'erreur ?

Réponse : Entre 4 et 5 %. Qu'allons-nous conclure ?

Réponse : Que nous devons examiner davantage de leucocytes pour que l'erreur devienne plus petite que le pourcentage observé.

18) — En examinant 136 choses, nous en trouvons 10 % qui appartiennent à une catégorie particulière. Quelle est, en gros, l'erreur, pour une probabilité de 9.999 pour 1 ?

Réponse : 10 %. (Trouvez la racine carrée du quotient de 13.621 divisé par 136.)

19) — En examinant 200 choses, nous en trouvons 80 qui appartiennent à une catégorie particulière. Quelle est l'erreur pour une probabilité de 99.999 pour 1 ?

Réponse : 15,3 %. (Divisez 46.785 par 200 et trouvez la racine carrée du quotient.)

20) — Travaillant avec une probabilité de 9.999 pour 1 et une erreur de 1 %, combien de choses devons nous examiner pour assurer un pourcentage de 41,7 % ?

Réponse : environ 36.789.

II. — FAÇON CORRECTE DE PROCEDER DANS UN TRAVAIL PRATIQUE.

Les tables peuvent être utilisées, en pratique, dans beaucoup de travaux sanitaires et médicaux, comme par exemple l'estimation de la fréquence de la mort ou de certains symptômes dans une maladie donnée ; ou de certains symptômes tels que l'augmentation du volume de la rate ou du rachitisme dans une population ou une formule leucocytaire chez un patient ou le nombre de colonies de microbes poussant sur culture.

Souvent, quoique pas toujours, les observateurs reconnaissent que l'échantillon doit être grand, mais habituellement ils fixent sa taille tout à fait arbitrairement. Ils disent par exemple que le nombre de leucocytes à examiner pour établir une formule leucocytaire est de 200 ou 500 ou 1000. Cette façon de procéder peut conduire à des résultats des plus erronés, parce que comme nous l'avons vu la taille de l'échantillon dépend de trois facteurs :

- 1° Le degré de certitude que nous voulons atteindre ;
- 2° Le pourcentage de l'erreur statistique ou le degré d'exactitude qui nous convient ;
- 3° Le pourcentage que nous avons tiré du rapport des choses à caractère particulier sur toutes celles que nous avons observées.

Le procédé correct sera donc :

- 1) — Il nous faut décider avec précision quel degré

de certitude nous voulons atteindre et le pourcentage d'erreur que nous admettons.

Ceci dépendra de l'importance de travail et du temps que nous avons à lui consacrer. Pour des recherches strictement scientifiques ou de grande portée sanitaire, nous devons rechercher un grand degré de probabilité et une marge d'erreur très petite : une probabilité de 99.999 pour 1 et une erreur de 1 %. Ce haut degré d'exactitude sera indispensable surtout quand nous voudrions établir des comparaisons de résultats obtenus à des moments différents, comme par exemple quand nous désirons savoir si le pourcentage de leucocytes mononucléaires augmente avec le progrès d'une maladie (voir exemples 13 et 16) ; ou si une épidémie décroît avec le temps. Dans ce cas, il est absolument indispensable que les erreurs statistiques obtenues aux deux moments différents ne soient pas assez grandes pour chevaucher. D'autre part nous pouvons souvent nous permettre d'adopter de plus faibles degrés de certitude et de plus hauts degrés de pourcentage d'erreur, quand nous cherchons simplement quelque évidence qui confirme une théorie, ou quand les différences entre deux évaluations successives sont si grandes ou si frappantes que même un grand pourcentage d'erreur ne peut égarer le jugement. Ici comme pour le chapitre suivant, nous devons être guidés souvent par les progrès de l'ouvrage. Mais aussitôt que nous nous décidons sur les trois points ci-dessus nous déterminerons quelle table et quelle colonne de cette table nous devons utiliser.

2) — Avant de nous fixer sur la taille de l'échantillon, nous devons nous efforcer d'obtenir, par un essai, une estimation grossière, du pourcentage que nous voulons déterminer.

Supposons par exemple, que nous ayons à faire une formule leucocytaire. Nous n'espérons pas, d'un côté, admettre une grande erreur statistique et nous ne voulons pas perdre trop de temps à examiner un échantillon trop grand. Supposons que la Table 9 pour 1 nous suffise. Commençons par examiner 100 leucocytes. Supposons que 40 % soient des mononucléaires. Nous pouvons alors chercher dans la table combien nous devons examiner de leucocytes pour donner à ce résultat, une certitude suffisante.

Si une erreur de 6 % semble suffire, nous savons que nous devons examiner 81 leucocytes de plus. Si nous ne voulons admettre que 2 % d'erreur, nous aurons à examiner 1.524 leucocytes de plus.

Au fur et à mesure que nous avançons maintenant dans notre formule leucocytaire, nous trouvons que le pourcentage des mononucléaires change considérablement. Nous avons examiné : 200, 300, 400 leucocytes de plus ; (nous ne devons pas calculer le pourcentage d'erreur pour chaque tranche successive de 100 leucocytes, mais pour le nombre total examiné depuis le début.) Finalement, quand environ 1.400 leucocytes ont été comptés, nous approchons de la limite prévue pour 2 % d'erreur. Supposons qu'à 1559 leucocytes, le pourcentage trouvé se maintienne à 36 % environ. concordant ainsi avec la table. Nous nous arrêtons alors, ayant obtenu un rapport de 36 % avec une probabilité de 9 contre 1 et une erreur statistique de 2 %. C'est-à-dire, la probabilité est de 9 chances contre une que la vérité se trouve comprise entre 38 % et 34 %. Nous n'avons pas perdu de temps pour atteindre ainsi ce résultat.

Si nous voulons obtenir de hauts degrés de probabi-

lité ou de faibles degrés d'erreurs, ou les deux ensemble, il faudra essayer une évaluation préliminaire assez grossière au moyen d'un petit échantillon de 100 leucocytes. Ce sera plus simple que de prendre pour cela du premier coup des échantillons de 500 ou 1.000 ou plus, suivant les cas.

3) — Naturellement dans les formules leucocytaires nous attendons souvent à l'avance un pourcentage déjà connu. Ainsi, les éosinophiles sont généralement en petit nombre et les polynucléaires en grand nombre. Dans ce cas, nous pouvons nous permettre de juger plus grossièrement la taille de l'échantillon.

4) — C'est une grave erreur de croire qu'un grand nombre d'échantillons compense un travail peu soigné. Les tables d'erreurs statistiques sont basées sur l'hypothèse que chaque chose examinée a été soigneusement assignée à sa vraie catégorie. Quand aux choses qui n'ont pas été certainement assignées à leur propre classe, elles doivent être rejetées, mais on doit déterminer leur proportion par rapport au nombre total de choses au moyen des évaluations de probabilité et d'erreurs, à l'aide précisément des mêmes méthodes que celles décrites plus haut. Elles constituent alors une troisième classe par elles-mêmes.

5) — Si les choses étudiées peuvent être divisées en trois classes ou plus, déterminons séparément la proportion de chaque classe à l'entier. Ceci ne demande pas de méthode particulière : nous établissons simplement les rapports en partant des chiffres collationnés, mais nous devons faire attention à ce que les échantillons soient suffisants par eux-mêmes pour chaque classe.

6) — Si nous examinons des échantillons successifs

(tels que des leucocytes) et que nous découvrons, dans chaque échantillon pris à la suite, que le pourcentage tend toujours vers une même direction, c'est-à-dire croît ou décroît, nous devons suspecter que quelque influence autre que le pur hasard préside à ces variations parallèles. Le nombre d'échantillons successifs que nous devons examiner pour confirmer cette hypothèse dépendra de la nature du matériel mais devra toujours être grand. Si le pourcentage observé croît toujours ou toujours diminue dans au moins cinq échantillons successifs, nous devons orienter nos recherches sur le point en question, ou en référer à un statisticien exercé.

7) — La probabilité ou degré de certitude qui est généralement accepté par les statisticiens comme traduisant la presque certitude est : 49.999 contre 1. On peut obtenir les chiffres correspondant à cette probabilité en multipliant les chiffres correspondant à une probabilité de 9999 contre 1 (Table A2) par le facteur 1,189 ; ou ce qui revient sensiblement au même en augmentant les chiffres de la table A2 de 20 %.

8) — Il faut toujours prendre grand soin de choisir les échantillons au hasard et de ne pas les choisir avec inclination consciente ou inconsciente.

III. — ECHANTILLONS PROPORTIONNELLEMENT GRANDS

Nous nous sommes occupés d'échantillons proportionnellement petits, c'est-à-dire d'échantillons qui sont petits, comparés au nombre total des choses que nous voulons examiner. Par exemple, si nous voulons étudier tous les individus d'une région ou tous les leucocytes du corps d'un malade, il nous sera ra-

rement possible d'en examiner plus qu'une assez petite proportion ; mais il y a des cas où le nombre total des choses n'est pas si grand que nous ne puissions en examiner une proportion plus considérable.

Supposons par exemple que nous désirions connaître la proportion d'enfants atteints de splénomégalie, non dans le monde entier, mais dans un grand village de 200 enfants. Supposons que nous n'ayions le temps d'en examiner que 50. Cet échantillon est proportionnellement grand. Il est le quart du nombre total des enfants du village. Si nous supposons qu'il nous est possible d'examiner 180 enfants, l'échantillon est si grand qu'il approche le nombre entier des éléments en études. Evidemment, en examinant ces grands échantillons, nous nous approchons de beaucoup plus près de la certitude absolue que nous ne le permet l'usage des Tables d'erreur, même à la probabilité de 99.999 contre 1.

Dans le deuxième cas par exemple, nous n'aurions à examiner que 20 enfants de plus pour obtenir une certitude absolue.

Nous comprenons que dans ces cas on doit corriger les Tables, les adapter à l'étude des grands échantillons. Pour cela, il suffit de multiplier l'erreur statistique donnée dans les tables par le facteur :

$$\sqrt{1 - \frac{n-1}{N-1}}$$

ou n est le nombre d'éléments que contient l'échantillon et N le nombre total des éléments observés.

En examinant ce « facteur correction », nous voyons que lorsque n est très petit par rapport à N , le terme devient si petit qu'il peut être négligé et le « facteur correction » devient : la racine carrée

de moins un sur moins un, c'est-à-dire un. Les tables restent alors tout à fait exactes.

...Quand $n = N$, c'est-à-dire quand l'échantillon comprend toutes les choses étudiées, le terme est égal à 1 et le facteur correction devient la racine carrée de $1 - 1$, c'est-à-dire 0 ; et quand le facteur correction multiplie les erreurs statistiques des Tables, il les efface. En d'autres mots, il n'y a plus d'erreurs statistiques, parce que nous avons la certitude absolue.

Entre ces valeurs, le facteur correction est une fraction ordinaire qui peut facilement être calculée. Ainsi si $n = 50$ et $N = 200$ le facteur est : la racine carrée de $150/199$, c'est-à-dire : 0,868. Ceci réduit l'erreur statistique, mais pas de beaucoup.

Si $n = 180$ le facteur est la racine carrée de $20/199$ soit : 0,317 et l'erreur statistique est réduite considérablement.

Si le nombre total des choses examinées est quelque peu considérable, le facteur correction devient presque le même que la racine carrée de $1 - n/N$.

Dans ce cas, nous pouvons donner une table des différentes valeurs du facteur correction, qui correspond aux différentes valeurs de n/N . Dans la table suivante, le rapport des choses examinées aux choses totales, (n/N) est donné en pourcentage et les valeurs correspondantes du facteur correction sont inscrites au-dessous.

TABLE B

100 $n/N =$	5 %	10 %	15 %	20 %	25 %
facteur =	0,97	0,95	0,92	0,89	0,87
100 $n/N =$	30 %	35 %	40 %	45 %	50 %
facteur =	0,84	0,81	0,77	0,74	0,71

100 n/N =	55 %	60 %	65 %	70 %	75 %
facteur =	0,67	0,63	0,54	0,55	0,50
100 n/N =	80 %	85 %	90 %	95 %	100 %
facteur =	0,45	0,39	0,32	0,22	0,00

Ainsi si nous examinons la moitié des éléments totaux, les erreurs statistiques sont réduites des 7/10 de leur valeur, donnée dans les tables A. On verra que ce facteur détermine une légère différence dans l'erreur statistique à moins que le nombre des choses examinées soit plus grand que le dixième total.

Ex. : 21) — 9 des 141 enfants d'un village sont absents. En examinant le reste, nous en trouvons 14 (10 %) qui ont une splénomégalie. Quelle est l'erreur statistique pour une probabilité de 9.999 contre un ?

Réponse : 2,5 %. Le facteur est 1/4 (voir aussi l'ex. 18.)

22) — Il y a 1.000 individus dans une île. Sur 100 on en trouve 60 atteints de filariose. Quelle est l'erreur statistique pour une probabilité de 9 contre 1 ?

Réponse : environ 7,7 %.

23) — Sur 884 individus, dans une ville, les 3/4 sont examinés et 73 sont malades. Quel est le pourcentage des malades dans toute la ville ?

Réponse : entre 9 et 13 % pour une probabilité de 999 pour 1.

IV. — LE NOMBRE DES ÉLÉMENTS D'UNE CLASSE DONNÉE EXISTANT DANS UNE SURFACE, UN VOLUME, OU UN TEMPS DONNÉS.

Supposons que nous voulions connaître le nombre de choses, contenues dans une surface, un volume ou un temps donnés. Le seul moyen d'obtenir une certitude est de compter tous les éléments. Si nous

n'avons pas le temps, nous ne compterons que le nombre de choses contenues dans un échantillon défini, volume, surface, ou temps et nous déduirons de ce nombre, le nombre probable de choses contenues dans le volume, la surface, ou le temps tout entier. Une question se pose alors : quelle doit être la valeur de l'échantillon pour réduire à un certain degré de probabilité l'erreur statistique à un pourcentage que nous aurons choisi ?

1° Exemple : Nous voulons savoir combien il y a de petites pierres dans un million de « pieds cubiques » de gravier. Nous ne pouvons les compter toutes et devons par conséquent nous contenter de déterminer le nombre de petites pierres contenues dans 1, 2, ou plusieurs « pieds cubiques ». Dans combien de pieds cubiques devront nous compter les pierres pour pouvoir calculer le total de toutes ces pierres avec le degré de précision que nous aurons choisi ?

2° Exemple : Nous voulons savoir combien il y a de leucocytes ou de parasites dans le sang total d'un malade. Dans quelle quantité de sang devons nous compter les éléments ?

Pour que notre résultat soit le plus exact possible, prendrons nous 1, 2, ou plusieurs mm³ comme échantillon ?

Tout d'abord, pour avoir le droit d'employer la méthode d'échantillonnage, nous devons être assurés que les éléments qui nous occupent sont également répartis dans la superficie, le volume ou le temps. Si ce n'est pas le cas, nous ne pourrons jamais savoir avec exactitude, par notre échantillon, le nombre de tous les éléments.

Il eût été utile d'essayer d'évaluer la population de la Grande Bretagne en recensant tous les habitants du dixième de la superficie de la campagne anglaise. Mais la population n'est pas également répartie, elle est au contraire rassemblée spécialement dans certains districts et villes en vertu de certaines lois économiques. De même, nous ne pouvons estimer, en prélevant des échantillons de sang périphériques, combien il y a de parasites sanguins dans le corps tout entier, à moins que nous ne sachions que ces organismes ne se rassemblent pas spécialement en certains points de la circulation, en périphérie, en ce qui concerne cet exemple...

Pour résoudre ces problèmes, la première méthode qui vient à l'esprit est la suivante : compter le nombre trouvé dans quelques échantillons et multiplier ce nombre par le nombre total d'échantillons contenus dans la superficie, le volume ou le temps tout entier.

Exemple :

Si les petites pierres sont également réparties dans un million de « pieds cubiques » de gravier, il y aura dans un million de « pieds cubiques » un million de fois le nombre de petites pierres qui existent dans un « pied cubique ».

Il n'en est rien.

Il est possible par hasard que les petites pierres soient particulièrement grandes dans le premier « pied cubique », pris comme échantillon. Elles sont alors assez peu nombreuses.

Il est également possible que, toujours par hasard, les trypanosomes dans un premier échantillon de sang, soient exceptionnellement nombreux ou exceptionnellement rares, suivant les cas.

Nous ferons alors une évaluation tout-à-fait erronée, si nous nous fions entièrement à la simple, mais infidèle méthode qui vient d'être mentionnée.

Pour obtenir des estimations exactes dans ces cas, il faudra, pour affirmer si les choses sont réellement réparties avec égalité, demander le secours d'un statisticien averti, et aussi faire une étude approfondie de la qualité du matériel considéré.

Pour les besoins de cet article qui ne vise que l'utilisation du calcul des probabilités par les statisticiens de la Médecine, nous donnons la table suivante qui sert à déterminer si les éléments sont également répartis dans l'aire, le volume, ou le temps.

TABLE C

<i>Probabilité</i>	<i>Erreurs statistiques</i>		
	1%	2%	3%
1 à 1	4.550	1.138	506
9 à 1	27.057	6.764	3.006
99 à 1	66.350	16.588	7.372
999 à 1	108.284	27.071	12.032
9999 à 1	151.338	37.835	16.815
99999 à 1	194.938	48.735	21.660

<i>Probabilité</i>	<i>Erreurs statistiques</i>			
	4%	5%	10%	20%
1 à 1	285	182	46	11
9 à 1	1.691	1.082	271	68
99 à 1	4.147	2.654	664	166
999 à 1	6.768	4.331	1.083	271
9999 à 1	9.459	6.053	1.513	378
99999 à 1	12.184	7.797	1.949	487

Supposons par exemple qu'on ait trouvé 3.006 petites pierres dans un « pied cubique » de gravier. Le nombre de pierres le plus probable dans un million de « pieds cubiques » du même gravier sera : 3.006 millions, avec une erreur de 3 % et une probabilité de 9 pour 1, ce qui veut dire que nous pouvons affirmer 9 contre 1 que le nombre total de pierres dans un million de pieds cubiques sera compris entre : 3.096 et 2.916 millions.

On a découvert 271 trypanosomes dans un mm³ de sang chez un malade pesant 60 kgs -74. Son corps entier contient 3 millions de mm³ de sang. Nous pouvons affirmer 9 contre 1 que le nombre le plus probable de trypanosomes dans tout son sang sera compris entre 894.300.000 et 731.700.000 (10 % d'erreur). Et si nous avons compté 4.200 leucocytes par mm³ de sang, nous pouvons dire 9.999 contre 1 que le sang total du malade contient entre 13.356 et 11.844 millions de leucocytes (6 % d'erreur).

Nous avons montré dans le chapitre I, quelle était la méthode pour trouver le nombre des éléments à compter pour des erreurs supérieures à 6 % : il suffit de diviser les chiffres inscrits sous le pourcentage 1 %, dans la table dont nous aurons choisi la probabilité, par le carré du pourcentage requis.

Ainsi, pour une erreur de 10, % nous divisons par 100 et pour une erreur de 31,6 % nous divisons par 1.000.

Exemple :

Si nous trouvons 27 croissants de malaria dans 1 mm³ de sang du même malade, nous devons affirmer 9 contre 1 que ce sang contient en totalité environ 56 à 106 millions de croissants (en admettant

toujours que ces croissants soient également distribués dans le sang).

La table ci-dessus ne sert qu'aux échantillons proportionnellement petits, comme il est dit plus haut... Quand l'échantillon est plus grand que le $1/10$ du matériel total, nous devons employer le facteur correction du Chapitre III, plutôt que la Table C. Si nous comptons 1.500 leucocytes dans $1/4$ de mm^3 de sang et que nous désirons savoir le nombre de leucocytes qui se trouvent dans 1 mm^3 , la table ci-dessus nous donne environ 10 % d'erreur, pour une probabilité de 9.999 pour 1 ; tandis que par la table B, où le facteur correction est de 0,87 quand $n/N = 1/4$ ou 25 %, l'erreur est non de 10 % mais de 8,7 %. Nous pouvons dire alors avec une probabilité de 9.999 contre 1 que le nombre de leucocytes contenus dans 1 mm^3 de sang est compris entre 6.522 et 5.478. Mais pour le total du sang, c'est-à-dire pour trois millions de mm^3 , l'erreur de 10 % doit être maintenue, parce que le quart de mm^3 est devenu maintenant un échantillon proportionnellement petit. Le nombre de leucocytes du corps entier sera compris entre 19.800.000 et 16.200.000.

Nous pouvons nous servir d'une autre façon de la Table C.

Supposons que nous ayons compté 664 éléments dans un échantillon l'erreur est de 10 % pour une probabilité de 99 contre 1, nous pouvons être sûrs à 99 chances contre une, que le nombre d'éléments dans les échantillons futurs de la même taille sera compris entre 730, 4 et 597,6 (en admettant que l'échantillon soit proportionnellement petit).

Cette méthode nous permet de ne pas déterminer

à l'avance, la taille de l'échantillon.

Dans les numérations sanguines par exemple, seul le nombre d'éléments au mm³ de sang, nous intéresse, et non pas le nombre d'éléments contenus dans le corps entier. Le plus souvent nous voulons savoir si le nombre augmente ou diminue.

Si le nombre d'éléments du second échantillon examiné se trouve en dehors des marges d'erreur trouvées dans le premier, nous pouvons admettre pour une probabilité appropriée, qu'il y a eu accroissement au diminution.....

Exemples :

24) — Nous avons compté 4.250 hématies dans 1 millième de mm³ de sang. Quel est le nombre d'hématies dans un mm³ pour une probabilité de 99 contre 1 ?

Réponse : Environ entre 4.420.000 et 4.080.000.

25) — Combien pouvons-nous trouver d'hématies dans un second échantillon de la même taille, pour une probabilité de 999 contre 1 ?

Réponse : Environ entre 4.463.000 et 4.037.000.

26) — Il y a une semaine, nous avons trouvé 490 hématies dans 1/10.000 de mm³ du sang d'un malade. Aujourd'hui nous n'en trouvons plus que 294 dans un échantillon de même valeur. Pouvons-nous évaluer avec une probabilité de 999.999 contre 1 qu'il y a eu diminution ?

Réponse : Non, car les erreurs chevauchent.

Pouvons-nous évaluer avec la probabilité de 9 à 1, qu'il y a eu diminution ?

Réponse : Oui, car les erreurs ne chevauchent pas.

27) — Du sang a été dilué 100 fois. Dans un cen-

tⁱème de mm³ du mélange, nous trouvons 500 hématies. Combien devons-nous trouver d'hématies dans 1 mm³ du même sang pour une probabilité de 999 contre 1 ?

Réponse : 5.000.000 avec une erreur de 14,7 %.

28) — Nous avons trouvé 553 éléments dans un échantillon quelconque. Nous prélevons 100 échantillons identiques. Dans combien l'erreur sera plus grande que 7 % ?

Réponse : Dans 10.

29) — Un personnage officiel récemment nommé trouve qu'il est obligé d'écrire 150 lettres durant la première semaine de son affectation. Combien de lettres avec une probabilité de 999 contre 1 devra-t-il écrire chaque semaine à venir ?

Réponse : 150 avec une erreur de 26,87 %.

30) — Nous avons trouvé un embryon de filaire dans 1 mm³ de sang. Que pouvons-nous conclure du nombre des embryons de filaire contenu dans les 3.000.000 de mm³ de la circulation entière.

Réponse : Les chances sont de 1 contre 1 que l'erreur soit moins grande que 67,5 %, et que le nombre total des embryons de filaire dans la circulation sont compris entre : 5.025.000 et 975.000....

TABLE A 1
99.999 contre 1

%	1%	2%	3%	4%	5%	6%
1	1930	483	215	121	78	54
2	3821	956	425	239	153	107
3	5673	1419	631	355	227	158
4	7486	1872	832	468	300	208
5	9260	2315	1029	579	371	258
6	10995	2749	1222	688	440	306
7	12691	3173	1410	794	508	353
8	14348	3587	1595	897	574	399
9	15968	3992	1774	998	639	444
10	17545	4387	1950	1097	702	488
11	19085	4772	2121	1193	764	531
12	20586	5147	2288	1287	824	572
13	22048	5512	2450	1378	882	613
14	23471	5868	2608	1467	939	652
15	24855	6214	2762	1554	995	691
16	26200	6550	2912	1638	1048	728
17	27506	6877	3057	1720	1101	765
18	28773	7194	3197	1799	1151	800
19	30000	7500	3334	1875	1200	834
20	31191	7798	3466	1950	1244	867
21	32340	8085	3594	2022	1294	899
22	33451	8363	3717	2091	1339	930
23	34524	8631	3836	2158	1381	959
24	35557	8889	3951	2223	1423	988
25	36551	9138	4062	2285	1463	1016
26	37506	9377	4168	2345	1501	1042
27	38423	9606	4270	2402	1537	1068
28	39300	9825	4367	2457	1572	1092
29	40138	10035	4460	2509	1606	1115
30	40936	10234	4549	2559	1638	1138
31	41698	10425	4634	2607	1668	1159
32	42419	10605	4714	2652	1697	1179
33	43101	10776	4789	2694	1724	1198
34	43744	10936	4861	2734	1750	1216
35	44349	11088	4928	2772	1774	1232
36	44914	11229	4991	2808	1797	1248
37	45440	11360	5049	2840	1818	1263
38	45928	11482	5104	2871	1837	1276
39	46366	11592	5152	2898	1855	1288
40	46785	11697	5199	2925	1872	1300
41	47156	11789	5240	2948	1887	1310
42	47487	11872	5277	2968	1900	1320
43	47780	11945	5309	2987	1912	1328
44	48033	12009	5337	3003	1922	1335
45	48247	12062	5361	3016	1930	1341
46	48423	12106	5381	3027	1937	1346
47	48559	12140	5396	3035	1943	1349
48	48657	12165	5407	3042	1947	1352
49	48715	12179	5413	3045	1949	1354
50	48735	12184	5415	3046	1950	1354

TABLE A 2
9.999 contre 1

%	1%	2%	3%	4%	5%	6%
1	1449	375	167	94	60	42
2	2967	742	330	186	119	83
3	4404	1101	490	276	177	123
4	5812	1453	646	364	233	162
5	7189	1798	799	450	288	200
6	8536	2134	949	534	342	238
7	9852	2463	1095	616	395	274
8	11139	2785	1238	697	446	310
9	12395	3099	1378	775	496	345
10	13621	3406	1514	852	545	379
11	14816	3704	1647	926	593	412
12	15982	3996	1776	999	640	444
13	17117	4280	1902	1070	685	476
14	18221	4556	2025	1139	729	507
15	19296	4824	2144	1206	772	536
16	20340	5085	2260	1272	814	565
17	21354	5339	2373	1335	855	594
18	22338	5585	2482	1397	894	621
19	23291	5823	2588	1456	932	647
20	24215	6054	2691	1514	969	673
21	25107	6277	2790	1570	1005	698
22	25970	6493	2886	1624	1039	722
23	26802	6701	2978	1676	1073	745
24	27605	6902	3068	1726	1105	767
25	28376	7094	3153	1774	1135	789
26	29118	7280	3226	1820	1165	809
27	29829	7458	3315	1865	1193	829
28	30510	7628	3390	1907	1221	848
29	31160	7790	3463	1948	1247	866
30	31781	7946	3532	1987	1272	883
31	32372	8093	3597	2024	1295	900
32	32932	8233	3660	2059	1317	915
33	33461	8366	3718	2092	1339	930
34	33960	8490	3774	2123	1359	944
35	34430	8608	3826	2152	1378	957
36	34868	8717	3875	2180	1395	969
37	35277	8820	3920	2205	1441	980
38	35655	8914	3962	2229	1427	991
39	36004	9001	4001	2250	1441	1000
40	36322	9081	4036	2270	1453	1009
41	36609	9153	4068	2289	1465	1017
42	36866	9217	4097	2305	1475	1025
43	37093	9274	4122	2319	1484	1031
44	37290	9323	4144	2331	1492	1036
45	37459	9365	4162	2342	1499	1041
46	37593	9399	4177	2350	1504	1045
47	37699	9425	4189	2357	1508	1048
48	37774	9444	4198	2361	1511	1050
49	37820	9455	4203	2364	1513	1051
50	37835	9459	4204	2365	1514	1051

TABLE A 3
999 contre 1

%	1%	2%	3%	4%	5%	6%
1	1072	268	120	67	43	30
2	2123	531	236	133	85	59
3	3152	788	351	197	127	88
4	4159	1040	463	260	167	116
5	5144	1286	572	322	206	143
6	6107	1527	679	382	245	170
7	7050	1763	784	441	282	196
8	7970	1993	886	499	319	222
9	8869	2218	986	555	355	247
10	9746	2437	1083	610	390	271
11	10601	2651	1178	663	424	295
12	11435	2859	1271	715	458	318
13	12247	3062	1361	766	490	341
14	13038	3260	1449	815	522	363
15	13807	3452	1535	863	553	384
16	14554	3639	1617	910	583	405
17	15279	3820	1698	955	611	425
18	15983	3996	1776	999	640	444
19	16665	4167	1852	1042	667	463
20	17326	4332	1925	1083	693	482
21	17965	4492	1995	1123	719	499
22	18582	4646	2065	1162	744	517
23	19178	4795	2131	1199	767	533
24	19751	4938	2195	1235	791	549
25	20304	5076	2256	1269	813	564
26	20834	5209	2315	1303	833	579
27	21343	5336	2372	1334	854	593
28	21830	5458	2426	1365	874	607
29	22296	5574	2478	1394	892	620
30	22740	5685	2527	1422	910	632
31	23162	5691	2574	1448	927	644
32	23562	5891	2618	1473	943	655
33	23942	5986	2661	1497	958	666
34	24299	6075	2700	1519	972	675
35	24635	6159	2738	1540	986	685
36	24949	6238	2773	1560	998	694
37	25241	6311	2805	1578	1010	702
38	25512	6378	2835	1595	1021	709
39	25761	6441	2863	1611	1031	716
40	25989	6498	2888	1625	1040	722
41	26194	6549	2911	1638	1048	728
42	26378	6595	2931	1649	1055	733
43	26541	6636	2949	1659	1062	738
44	26682	6671	2965	1668	1068	742
45	26800	6700	2978	1675	1072	745
46	26898	6725	2989	1682	1076	748
47	26974	6744	2997	1686	1079	750
48	27028	6757	3004	1689	1081	751
49	27061	6766	3007	1692	1082	751
50	27071	6768	3008	1692	1083	752

TABLE A 4
99 contre 1

%	1%	2%	3%	4%	5%	6%
1	658	165	674	42	27	19
2	1301	325	145	82	52	37
3	1931	483	215	121	78	54
4	2549	638	284	160	102	71
5	3151	788	350	197	126	88
6	3743	936	416	234	150	104
7	4319	1080	480	270	173	120
8	4884	1221	543	306	196	136
9	5434	1359	604	340	218	151
10	5973	1494	664	374	239	166
11	6496	1624	728	406	260	181
12	7007	1752	779	438	281	195
13	7504	1876	834	469	301	209
14	7989	1998	888	500	320	222
15	8461	2116	940	529	339	236
16	8918	2230	991	558	357	248
17	9361	2341	1040	586	375	260
18	9794	2449	1089	613	392	273
19	10211	2553	1135	639	409	284
20	10616	2654	1180	664	425	295
21	11007	2752	1223	688	441	306
22	11386	2847	1265	712	456	317
23	11752	2938	1306	735	470	327
24	12104	3026	1345	757	485	337
25	12441	3111	1383	778	498	346
26	12767	3192	1419	798	511	355
27	13078	3270	1453	818	523	364
28	13376	3344	1487	836	535	372
29	13682	3421	1521	856	548	381
30	13934	3484	1549	871	558	388
31	14193	3549	1577	888	568	395
32	14439	3610	1605	903	578	402
33	14672	3668	1631	917	587	408
34	14890	3723	1655	931	596	414
35	15096	3774	1678	944	604	420
36	15287	3822	1699	956	612	425
37	15467	3867	1719	967	619	430
38	15633	3909	1737	978	625	435
39	15787	3947	1755	987	632	439
40	15925	3982	1770	996	637	443
41	16051	4013	1784	1004	642	446
42	16164	4041	1796	1011	647	449
43	16263	4066	1807	1017	651	452
44	16349	4088	1817	1022	654	455
45	16422	4106	1825	1027	657	457
46	16482	4121	1832	1031	660	458
47	16530	4133	1837	1034	662	460
48	16562	4141	1841	1036	663	461
49	16582	4146	1843	1037	664	461
50	16588	4147	1844	1037	664	461

TABLE A 5

9 contre 1

%	1%	2%	3%	4%	5%	6%
1	268	67	30	17	11	8
2	531	133	59	34	22	15
3	788	197	88	50	32	22
4	1039	260	116	65	42	29
5	1286	322	143	81	52	36
6	1526	382	170	96	62	43
7	1762	441	196	111	71	49
8	1992	498	222	125	80	56
9	2216	554	247	139	89	62
10	2436	609	271	153	98	68
11	2649	663	295	166	106	74
12	2858	715	318	179	115	80
13	3061	766	341	192	123	86
14	3258	815	362	204	131	91
15	3450	863	384	216	139	96
16	3637	910	405	228	146	102
17	3818	955	525	239	153	107
18	3994	999	444	250	160	111
19	4164	1041	463	261	167	116
20	4330	1083	482	271	174	121
21	4489	1123	499	281	180	125
22	4643	1161	516	291	186	129
23	4792	1198	533	300	192	134
24	4936	1234	549	309	198	138
25	5074	1269	564	318	203	141
26	5206	1302	579	326	209	145
27	5333	1334	593	334	214	149
28	5455	1364	607	341	219	152
29	5571	1393	619	349	223	155
30	5682	1421	632	356	228	158
31	5788	1447	644	362	232	161
32	5888	1472	655	368	236	164
33	5983	1496	665	374	240	167
34	6072	1518	675	380	243	169
35	6156	1539	684	385	246	171
36	6234	1559	693	390	250	174
37	6307	1577	701	395	253	176
38	6375	1594	709	399	255	178
39	6437	1610	716	403	258	179
40	6494	1624	722	406	260	181
41	6545	1637	728	410	262	182
42	6591	1648	733	412	264	184
43	6632	1658	737	415	266	185
44	6667	1667	741	417	267	186
45	6697	1675	745	419	268	187
46	6721	1681	747	421	269	187
47	6740	1685	749	422	270	188
48	6754	1689	751	423	271	188
49	6762	1691	752	423	271	188
50	6764	1691	752	423	271	188

TABLE A 6

1 contre 1

%	1%	2%	3%	4%	5%	6%
1	45	11	5	3	2	1
2	89	22	10	6	4	2
3	132	33	15	8	5	4
4	175	44	19	11	7	5
5	216	54	24	14	9	6
6	257	64	29	16	10	7
7	296	74	33	19	12	8
8	335	84	37	21	13	9
9	373	93	41	23	15	10
10	409	102	45	26	16	11
11	445	111	49	28	18	12
12	480	120	53	30	19	13
13	515	129	57	32	21	14
14	548	137	61	34	22	15
15	580	145	64	36	23	16
16	611	153	68	38	25	17
17	642	160	71	40	26	18
18	671	168	75	42	27	19
19	700	175	78	44	28	19
20	728	182	81	45	29	20
21	755	186	84	47	30	21
22	781	195	87	48	31	22
23	806	201	90	50	32	22
24	830	207	92	52	33	23
25	853	213	95	53	34	24
26	875	219	97	55	35	24
27	897	224	100	56	36	25
28	917	229	102	57	37	25
29	937	234	104	59	37	26
30	955	239	106	60	38	27
31	973	243	108	61	39	27
32	990	247	110	62	40	28
33	1006	251	112	63	40	28
34	1021	255	113	64	41	28
35	1035	259	115	65	41	29
36	1048	262	116	66	41	29
37	1060	265	118	66	42	29
38	1072	268	119	67	42	30
39	1082	271	120	68	43	30
40	1092	273	121	68	43	30
41	1100	275	122	69	44	31
42	1108	277	123	69	44	31
43	1115	279	124	70	44	31
44	1121	280	125	70	45	31
45	1126	281	125	70	45	31
46	1130	282	126	71	45	31
47	1133	283	126	71	45	32
48	1135	284	126	71	45	32
49	1137	284	126	71	45	32
50	1137	284	126	71	45	32

B

LE CALCUL DE L'ECART PROBABLE
PAR L'EMPLOI DES NOMOGRAMMES

II. — LES NOMOGRAMMES

Nous connaissons maintenant les Tables de Ross et la façon dont nous devons nous en servir.

Cependant, nous avons remarqué que ces auteurs, tout en signalant que les statisticiens admettent couramment une probabilité de 49.999 pour 1, soit 0,99998, comme représentant la presque certitude, ne donnent aucune Table spéciale pour cette probabilité. Ils indiquent bien une correction à faire subir aux chiffres de la Table A2, mais c'est tout.

Il est bizarre que la probabilité la plus souvent admise ne soit pas donnée sous forme de Table.

Nous aurions pu en dresser une en modifiant tous les chiffres de la Table A 2 de Ross et Stott. Il nous a paru plus intéressant de montrer pour cette probabilité de 0,99998, comment on se sert du Nomogramme du type à points alignés, imaginé en 1884 par M. M. d'Ocagne, membre de l'Académie des Sciences, Inspecteur Général des Ponts et Chaussées, Professeur à l'Ecole Polytechnique.

Construit par Stanislas Milet, Capitaine de Corvette, pour l'application au théorème de J. Bernouilli qui suppose une probabilité constante, le nomogramme que nous reproduisons a paru dans un article des archives des Instituts Pasteur de l'Afrique du Nord, Tome II, 1922 (p. 587).

Ce Nomogramme est double.

Le grand n'est construit que pour une valeur de $x = 3$ constante.

La formule de Laplace :

$$1 = \frac{x \sqrt{2pq}}{\sqrt{m}} \quad (\text{Formule n}^\circ 30)$$

donnant à la probabilité Θ x la valeur constante de 0,99998 la plus usuelle.

Le petit Nomogramme qui figure au bas et à droite du Grand, sert à transformer l'écart maximum conventionnel, quand on adopte pour x une autre valeur que 3, c'est-à-dire quand la probabilité varie de 0,999999985 à 0,5. Ces deux nomogrammes réunissent donc les avantages des six tables de Ross et Stott :

Le petit remplace les Tables à probabilité variable. Le grand remplace la Table A 2 modifiée pour la probabilité fixe et constante de 0,99998.

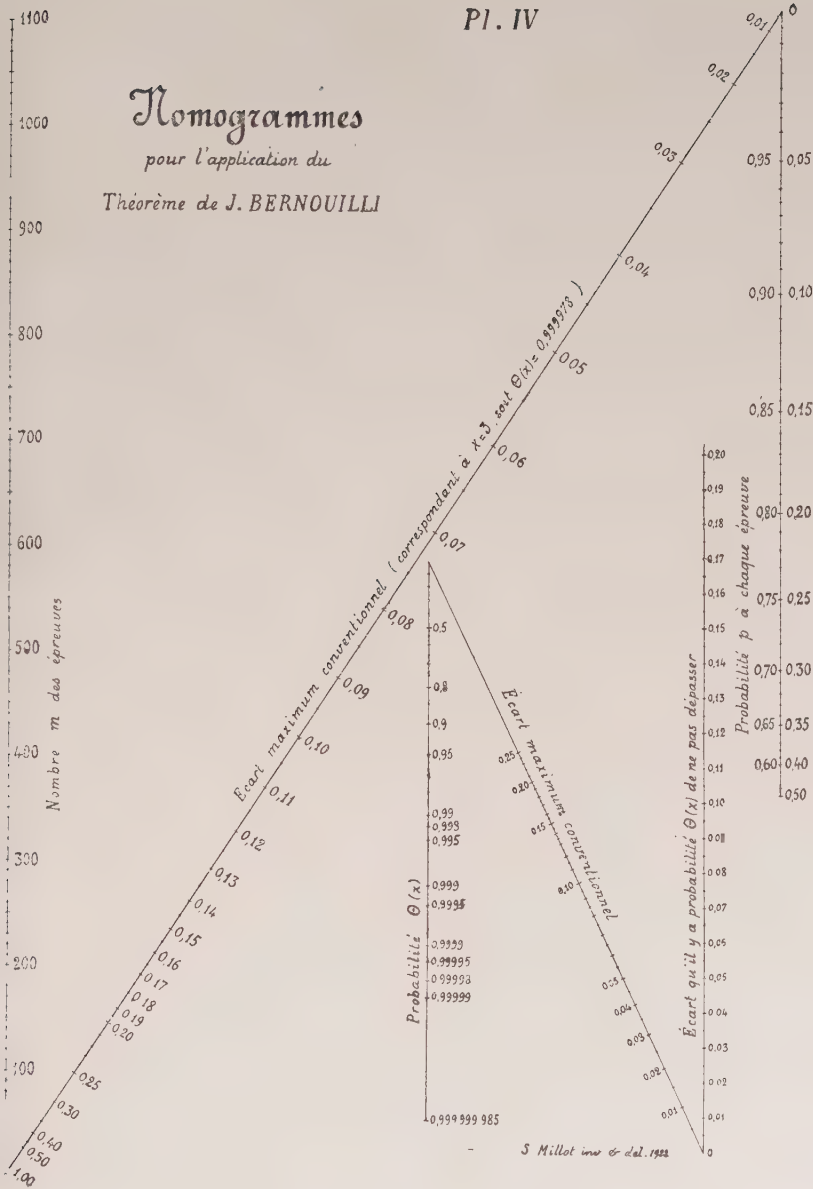
Leur mode d'emploi est simple. S. Millot dans son article nous en donne des exemples.

Pl. IV

Nomogrammes

pour l'application du

Théorème de J. BERNOULLI



PREMIER EXEMPLE DE CACUL SIMPLIFIE

« Supposons par exemple, un mélange très important de granulés contenant, lès uns un médicament très actif, les autres une matière inerte, les premiers formant les deux centièmes du mélange.

Si, après brassage, on prend 400 granules au hasard, on n'y trouvera pas obligatoirement les 8 granules actifs qui correspondraient exactement à la probabilité $p = 0,02$. Pour trouver l'écart maximum, on peut se servir du plus grand des deux nomogrammes de la planche IV, qui sont du type à points alignés, imaginés par M. Maurice d'Ocagne.

Joignant, par une ligne droite (un fil tendu par exemple) la graduation 0,02 de l'échelle de droite (probabilité p) à la graduation 400 de l'échelle de gauche (nombre m des épreuves), on lit, à la rencontre du fil avec l'échelle oblique du milieu, le nombre 0,03 qui est l'écart maximum conventionnel.

Dans l'échantillon de 400, la proportion des granules actifs pourra donc varier de $p - 0,03$ nombre négatif, auquel on substituera zéro, à $p + 0,03$, c'est-à-dire 0,05. Au lieu de 8 granules actifs on aura de zéro à 20 granules.

Si la proportion des granules actifs était dix fois plus forte dans le mélange, p vaudrait 0,20 et, dans l'espoir d'obtenir la même dose moyenne de huit granules actifs, on prendrait 40 granules au hasard dans le mélange. Le calcul, fait sur le nomogramme avec 40, à gauche (m) et 0,20 à droite (p) ferait trouver comme écart maximum 0,24, si bien que le nombre des granules actifs dans le groupe de 40 pourra varier entre zéro et le produit de 0,44 par 40, soit 17,6 ou 18.

Si, au lieu de granules on avait des grains 100 fois plus petits, il faudrait en prendre 40.000 dans le premier cas et 40.000 dans le second.

Le calcul pourra encore être fait avec les nombres 400 et 40, 100 fois plus petits, mais les erreurs devraient être divisées par 10 et seraient 0.003 dans le premier cas, et 0,024 dans le second. Il en résulterait un nombre de grains actifs compris entre 680 et 920, dans le premier cas et entre 704 et 896 dans le second cas.

Cet exemple semble prouver, bien que les résultats obtenus sur des échantillons de 40 à 400 ne puissent être très exacts, les deux propositions suivantes :

- 1° Quand on mélange des médicaments d'activité différente et fractionne ensuite le mélange, l'exactitude du dosage de l'élément énergétique est d'autant plus douteuse dans chaque fraction que la proportion des matières inertes est plus forte dans le mélange.
- 2° Quand un mélange de poudres médicamenteuses doit être fractionné, les poudres les plus fines sont celles qui, bien brassées, doivent donner les plus faibles écarts de dosage.

Tout cela suppose naturellement des granules ou des grains parfaitement égaux comme taille et poids.

Dans la pratique, il n'en est pas ainsi et les inégalités de dosage doivent être plus forte que ne l'indiquent les exemples théoriques.

42) — Il reste peu de choses à ajouter pour montrer le mode d'emploi des nomogrammes de la planche IV. Nous avons déjà vu que l'on peut ramener

les très grandes valeurs de m dans les limites de l'échelle de gauche en les divisant par exemple par 100 ce qui permet d'aller au delà de 100.000 ; l'écart trouvé doit alors être divisé par 10. Si l'on employait sur l'abaque le $1/4$ du nombre m , l'écart trouvé devrait être divisé par 2 racine carrée de 4 etc...

L'examen de la formule 30 montre que l'écart ne change pas de valeur quand on permute p et q , c'est-à-dire quand on remplace p par $.1 - p$).

Une même graduation de l'échelle de droite conviendra donc pour les probabilités de 0,10 et 0,90 ou 0,20 et 0,80... etc... C'est pourquoi cette échelle porte des chiffres sur ses deux côtés.

Le petit nomogramme en forme d' N qui figure au bas et à droite de la plauche IV, sert à transformer l'écart maximum conventionnel quand on adopte pour x une autre valeur que 3. Soit par exemple l'écart maximum conventionnel 0,05.

Si l'on veut un degré plus grand de certitude, tel que 0,999999985 correspondant pour x à la valeur 4, on fait passer une ligne droite (ou fil tendu) par la graduation 0,999999985 de l'échelle de gauche et par la graduation 0,05 de l'échelle oblique (écart maximum) et, dans le même alignement, on lit 0,038 sur l'échelle de droite. La probabilité est alors 0,999 pour que l'écart 0,038 ne soit pas dépassé.

Le petit nomogramme permet aussi facilement de passer d'un écart et de la probabilité correspondante à l'écart maximum conventionnel, ou bien, connaissant ce dernier écart, de déterminer la probabilité qu'un autre écart donné ne sera pas dépassé. L'écart maximum conventionnel étant par exemple 0,18, l'écart 0,07 correspondra à la probabilité 0,90... etc....

DEUXIEME EXEMPLE D'APPLICATION DU CALCUL SIMPLIFIE

« Sur les remarques du paragraphe précédent, on a basé une méthode d'appréciation de la valeur de diverses expériences comparatives, faites journellement dans les laboratoires.

Cette méthode consiste, lorsque, dans une expérience, la probabilité apparente est supérieure à 0,5, tandis que, dans une contre-expérience, elle est inférieure à 0,5, à comparer avec la probabilité 0,5 les deux zones de probabilité qui correspondent à l'expérience et à la contre-expérience.

L'emploi du nomogramme spécial rend alors le calcul extrêmement facile. »

Ce Nomogramme qui a été publié dans « La théorie nouvelle de la probabilité des causes », de S. Milot en 1925 chez Gauthier Villars et Cie, est comme celui que nous reproduisons constitué par deux échelles rectilignes et parallèles, portant des graduations équidistantes, relatives les unes, à droite, au nombre d'arrivées de l'événement considéré (a) de 0 à 100. Les autres à gauche sont relatives au degré de certitude $\Theta \times$ de 0,999999999998 à 0 ; puis de 0 à 0,999999999998.

Une troisième échelle, en arc hyperbole, part du zéro de l'échelle droite, relative au nombre d'arrivées de l'événement considéré, et a pour asymptote la deuxième échelle rectiligne (Θx).

Cet arc hyperbole, gradué d'après les valeurs de m (nombre d'épreuves) est spécial à une certaine valeur de Y (probabilité donnée), et si l'on veut envisager plusieurs valeurs de Y, il faut tracer sur le nomogramme plusieurs arcs, distincts, d'hyperbole.

Celui qui se trouve dans l'ouvrage précité est construit pour $Y = 1/2$. C'est un nomogramme donnant le degré de certitude que la probabilité supposée fixe et réelle d'un événement n'est pas $1/2$.

Nous renvoyons les lecteurs à l'ouvrage original, s'ils ont besoin d'utiliser ce nomogramme.

Nous donnons, écrit par S. Millot, un exemple de calcul simplifié, grâce à ce nomogramme.

« A l'Institut Pasteur d'Alger, MM. G. Senevet et P. Witas, ont essayé une réaction au sublimé-sulfate de soude sur des malades atteints de suppuration et sur d'autres non suppurants.

- I. — Sur 46 des suppurants, on a obtenu 40 réactions positives.
- II. — Sur 32 non suppurants (albuminuriques non compris), il n'y eut que deux réactions positives.
- III. — Sur 5 albuminuriques, il y eut 5 réactions positives.

Comparons avec 0,5 les zones de probabilité correspondant à ces trois séries d'observations.

$$\text{La formule (1) : } X = \text{val. abs. } \frac{m - 2a}{\sqrt{2m}}$$

nous donnera respectivement pour elles :

$$\begin{array}{lll} X = 3,54475 & \text{d'où } \Theta(X) = & 0,999999464 \\ = 3,50 & » & = 0,999999257 \\ = 1,581 & » & = 0,97464 \end{array}$$

Le nomogramme correspondant à cette formule nous donnerait encore plus rapidement ces valeurs de

$\Theta(X)$, d'où nous déduisons P par la formule :

$$p = 1/2 [1 + \Theta(X)]$$

Nous pouvons alors dire :

Il y a un degré de certitude.

- 1° — 0,999999732 pour que les réactions positives soient plus probables que les réactions négatives dans le cas de malades suppurants ;
- 2° — 0,999999629 pour que les réactions positives soient moins probables que les réactions négatives dans le cas de malades non suppurants (albuminuriques exceptés) ;
- 3° — 0,98732 pour que les réactions positives soient plus probables que les réactions négatives dans le cas d'albuminuriques.

Comparant les deux premiers cas, nous avons un degré de certitude minimum de 0,999999629 que ces deux cas correspondent, à l'égard de la réaction étudiée, à des états distincts, avec prédominance réelle des résultats positifs chez les suppurants.

Le calcul par la formule : (2)

$$Y = \frac{1}{m - m'} \left[(d - d') \pm \left(d \sqrt{\frac{m'}{m}} - d' \sqrt{\frac{m}{m'}} \right) \right]$$

aurait donné plus exactement, mais plus péniblement, pour le degré de certitude 0,999999680 ; il n'est pas nécessaire d'y recourir. La comparaison des 2 derniers cas ne donne qu'un degré de certitude minimum 0,98732 que ces 2 cas correspondent, à l'égard de la réaction étudiée, à des états distincts, avec prédominance réelle des résultats positifs chez les albuminuriques. Comme l'écart avec le maximum 0,999999629 est très grand, il serait plus indiqué que précédemment de recourir à la formule ci-dessus qui conduirait au de-

gré de certitude 0,999313, d'ailleurs insuffisant pour rendre la comparaison convainquante.

Si l'on avait disposé de nomogrammes construits pour d'autres valeurs de Y, on aurait pu les essayer pour resserer progressivement l'intervalle dans lequel est compris le degré de certitude....

Quoiqu'il en soit, le calcul des probabilités montre que, malgré la concordance des résultats donnés par les albuminuriques, il ne suffisait pas d'en examiner 5 pour établir scientifiquement une loi les concernant.

Continuée, l'expérience a fait obtenir, au total, 21 réactions positives sur 21 albuminuriques et la comparaison avec 1/2 donne :

$$X = 3,24, \Theta(X) = 0,9999954, P = 0,9999977$$

Le degré de certitude est devenu plus satisfaisant.

Si cette méthode de calcul simplifié, par comparaison avec 1/2, était adoptée, on pourrait examiner séparément chaque série d'observations, lui appliquer les formules (1) et (2) et noter le degré de certitude correspondant. Si ce degré de certitude semblait insuffisant, on s'efforcerait de prolonger la série d'observations jusqu'à ce que la valeur de P ait atteint le minimum que l'on se serait fixé.

Après avoir étudié, une à une, les diverses séries correspondant à des cas différents, on rangerait, les résultats en deux classes selon que la probabilité serait supérieure ou inférieure à 1/2, et l'on pourrait ensuite comparer l'une quelconque des séries de la première classe à l'une quelconque des séries de la deuxième, comme il a été fait ci-dessus pour les séries où les réactions positives prédominent et celles où, au contraire, elles sont en minorité.

La faculté que donne cette méthode d'évaluer isolément, la force probante d'une série d'épreuves, sans que l'on ait à s'inquiéter d'avance de l'usage que l'on pourra faire des résultats de cette série d'épreuves, pour les comparer avec d'autres, c'est évidemment une faculté précieuse qui, jointe à la simplicité du calcul fera sans doute agréer favorablement les procédés exposés au présent paragraphe. »

CHAPITRE IV

LA NOTION D'ECART PROBABLE MONTRE
L'INANITE DES DECIMALES DANS LES
POURCENTAGES PLUS GRANDS QUE L'UNITE

LA NOTION D'ECART PROBABLE MONTRE L'INANITE DES DECIMALES DANS LES POURCENTAGES PLUS GRANDS QUE L'UNITE

Il est fréquent de constater la présence de une ou plusieurs décimales dans les chiffres de pourcentage donnés par certains auteurs. La première décimale apprécie jusqu'au dixième, la deuxième jusqu'au centième.

Il est curieux de voir des appréciations qui veulent être si précises, s'établir en partant de méthodes qui en elles-mêmes comportent des marges d'erreur, non pas de l'ordre du dixième ni du centième, mais de l'ordre de l'unité, voir même de dizaines d'unités.

Les travaux de Ross et Stott, ceux de S. Millot, démontrent facilement l'inanité des approximations au dixième ou au centième, en appliquant, les premiers les « Tables of statistical error », le second, le calcul des probabilités, grâce au théorème de Jacques Bernouilli.

Il est un fait, c'est que les praticiens n'ont généralement pas la notion de la marge probable d'erreur de leurs calculs, et, poussés par une tendance légitime de la précision, ils ajoutent un zéro au dividende, une virgule au quotient, et arrivent ainsi à un pourcentage nanti d'une ou plusieurs décimales.

Il existe 2 cas bien distincts que nous étudierons successivement :

- 1° Les cas où le pourcentage est plus grand que l'unité ;
- 2° Le cas où il est plus petit que un.

I. LES DECIMALES DANS LES POURCENTAGES PLUS GRANDE QUE UN.

Nous prendrons comme exemple la publication de M. Sabrié (archives des Instituts Pasteur de l'Afrique du Nord, Tome II, 1922, p. 336). Cet auteur a observé 90 cuti réactions positives sur 198 sujets, soit :

$$\frac{90}{198} \cdot 100 = 45,45 \, \%$$

Ce qui signifie que sur 100 sujets pris comme échantillon-type de toute la population où a opéré M. Sabrié, quarante cinq sujets et presque un demi sujet (exactement quarante cinq centièmes de sujet), ont des cuti réactions positives. Ceci est évidemment un résultat d'une approximation très poussée, mais malheureusement illusoire, d'autant plus dangereux, qu'il peut en imposer à des lecteurs non avertis.

Si l'on cherche en effet à apprécier l'erreur probable par la méthode de Ross et Stott, on voit que pour les chiffres indiqués 90 sur 198 il y a :

- une chance sur 2 pour que l'erreur ne soit pas supérieure à 2 % ;
- une chance sur 10 pour que l'erreur ne soit pas supérieure à 6 %.

Si l'on veut serrer les faits de plus près et acquérir une certitude vraiment absolue :

99.999 sur 100.000 par exemple, on s'aperçoit qu'il

aurait fallu examiner 1341 sujets au lieu de 198 et que l'erreur probable aurait été de 6 % en plus ou en moins du chiffre de 45 % indiqué ; c'est-à-dire que le pourcentage certain aurait été compris entre 39 % et 51 %.

D'autre part, en utilisant le nomogramme d'application du théorème de Bernouilli (planche IV), nous voyons que pour les chiffres de 90 cuti réactions positives sur 198 sujets observés, l'erreur probable serait de plus ou moins 15 %, ce qui revient à dire que le pourcentage compté à partir de 198 cas serait compris entre 30 et 60 %.

On voit par cette critique ce que deviennent les décimales, alors que le chiffre des unités est grevé d'une marge d'erreur de 30 unités, c'est-à-dire de 300 dixièmes et de 3000 centièmes.

II. LES DECIMALES DANS LES POURCENTAGES PLUS PETITS QUE UN

Les décimales se justifient au contraire quand le chiffre est inférieur à l'unité, car, dans ce cas, la première décimale égale dix fractions de l'unité absente, la seconde 100 fractions, et l'appréciation comparative ne pourra se faire que sur elles. Ainsi 0,5 et 0,05 sont des chiffres 10 fois différents comme 50 et 5 ; et un pourcentage de 0,5 % correctement établi est 10 fois plus fort qu'un pourcentage de 0,05 %.

Il y a entre les deux pourcentages autant d'écart qu'entre 50 % et 5 %. La faute ne deviendrait comparable que si les approximations décimales étaient poussées au delà de chiffres supérieurs à la marge d'erreur probable que comporte le dénombrement des cas considérés.

On comprendra encore mieux la nécessité des décimales, quand le pourcentage est inférieur à l'unité en disant que s'il existe des valeurs intermédiaires sans décimales entre un et 10 il n'en existe pas entre un et zéro.

Enlever les décimales dans les pourcentages inférieurs à l'unité reviendraient à dire qu'entre 1 % et 0 % il n'y a pas d'intermédiaire.

CHAPITRE V

LE RÔLE DU HASARD DANS
L'INTERPRÉTATION DES CONCOMITANCES

IL EST INDISPENSABLE DE CONNAITRE
LE ROLE DU HASARD, AVANT D'INTERPRETER
LES CONCOMITANCES

Le « Post hoc, ergo propter hoc » peut être confirmé ou infirmé par :

- Le calcul des probabilités ;
- La méthode des statistiques témoins hétérogènes ;
- La méthode purement expérimentale.

Nous avons vu tout l'intérêt que présente la connaissance du « facteur hasard » dans les marges d'erreur obligatoire que comportent les statistiques médicales.

Etudions maintenant le rôle du hasard dans le problème des concomitances.

L'esprit humain est ainsi fait, qu'il veut toujours trouver les raisons de la concomitance de phénomènes et que le plus souvent il est tenté d'attribuer à leur concordance des rapports de cause à effet. Ce but louable qui a présidé à toutes les découvertes de génie, devient dangereux quand les concomitances sont dues seulement au hasard ; car la recherche des causes dans ce cas, demeure stérile et aiguille les chercheurs dans les dédales forts complexes de l'expérimentation vaine.

En médecine, les coïncidences ont de tout temps été à l'origine de légendes et de théories les plus fantaisistes. Malheureusement certaines d'entr'elles se sont

implantées dans les esprits et ce précisément par la méconnaissance des lois du hasard.

Le type en est la croyance de l'existence des maisons à cancers, et de l'hérédité cancéreuse, basée, la première sur la fréquence des néoplasmes dans certaines maisons, la deuxième sur la répétition des cas de tumeurs malignes dans une même famille.

Cette notion de maison à cancer a pris naissance à la suite des travaux de Fiessinger, d'Oyonnax, qui signala une sorte d'épidémie de cancérose dans un groupe de trois maisons où quatre concéreux succombèrent en quatre années. Dans sa thèse de Lyon 1892, Favre cite plusieurs cas analogues « de la contagion du cancer ». Gueliot, de Reims, Foucault, et d'autres... M. Chaton plus récemment résume dix observations des Docteurs Chancenotte, Courtois, Coulon, Dejeux, Guenot, Munier, Notton et Roudel qui en Franche-Comté seulement notèrent dans certains immeubles deux, trois et quatre décès par néoplasie, l'observation la plus édifiante relatant cinq décès par cancer dans une même maison en moins de quinze années.

Le Professeur Chaton, dans le *Bulletin de l'Académie de Médecine*, Tome 107, 1932, p. 875, sous le titre « de l'intérêt très particulier de l'étude des maisons à cancer et de la manière dont on pourrait y procéder », nous dit :

« Si dans une maison, un observateur médical constatait une proportion anormale de cas de diphtérie ou de fièvre typhoïde, il étudierait l'hygiène générale des lieux, il rechercherait une cause commune déjà connue, un porteur de germe, une contamination hydrique. »

Il est évident que la remarque du Professeur Chaton est très logique. Cependant avant de l'appliquer aux maisons « dites à cancers », il serait utile de démontrer que la fréquence des tumeurs malignes dans certains immeubles correspond à une fréquence anormal de néoplasie et pas seulement à de pures coïncidences déterminées par les lois du hasard.

CRITIQUE

I. PAR LE CACUL DES PROBABILITES

(emprunté à l'Avenir Médical xxx^e année, n° 2, Février 1933, Auguste Lumière et Paul Vignes.)

« Dans son exposé, M. Chaton ne fait mention ni du travail de Besson sur les maisons à Cancer, (1) ni des arguments que l'un de nous a opposés aux déductions de ce dernier auteur (2).

Besson a basé son étude sur un document statistique tiré du rapport de M. Juillerat au Préfet de la Seine et concernant le casier sanitaire des maisons de la Ville de Paris. Cet état démographique enregistré dans la capitale 24.957 décès par cancer de 1906 à 1914, décès dont la répartition, dans les immeubles de la ville s'est opérée de la façon suivante :

- 61.726 maisons sans cancéreux ;
- 14.784 maisons ayant chacune un décès par cancer ;
- 3.175 avec deux décès pour la même cause ;
- 708 avec trois décès ;

(1) BESSON. — « Le cancer et les maisons ». Annales d'hygiène publique, industrielle et sociale, Tome I, 1923 (p. 23).

(2) AUGUSTE LUMIÈRE. — « La contagion, l'hérédité et le hasard dans le cancer ». Bulletin de l'Académie de Médecine, Tome I, 1027 (p. 89).

181 avec quatre ;
 48 avec cinq ;
 13 avec six ;
 2 avec sept ;
 1 avec huit ;
 1 avec neuf.

Besson a alors cherché à établir, par le calcul des probabilités comment seraient répartis ces décès dans les 80.639 immeubles de Paris, en admettant que le hasard procède seul à ce classement. La comparaison des nombres donnés par ce calcul avec ceux qui ressortent de la statistique est indiquée dans le tableau I.

TABLEAU I

Nombre de décès par maison	0	1	2	3	4	5	6	7	8	9
Calcul des probabilités ...	59175	18414	2834	292	23	1				
Réalité	61726	14784	3175	708	184	48	13	2	1	1

La concordance qui est approchée pour 0, 1 et 2 décès par immeuble n'existe plus à partir de trois décès par maison et l'on pourrait croire à priori que des causes étrangères interviennent pour troubler la distribution ; mais Besson fait lui-même observer qu'il a introduit, dans le calcul une hypothèse toute gratuite, en admettant que tous les immeubles abritent le même nombre d'habitants supposé égal à 33, ce qui est fort éloigné de la vérité, les maisons comptant des nombres éminemment variables d'occupants, depuis quelques unités, jusqu'à la centaine, et plus encore.

Pour se faire une idée de l'influence du nombre d'habitants d'un immeuble sur la répartition théorique des cas de cancer, l'auteur a procédé à un second calcul, dans lequel il a divisé les maisons en deux groupes : l'un de 50.000 immeubles avec 28 habitants et l'autre de 30.639 immeubles comportant chacun 48 occupants.

Dans ces conditions la répartition est modifiée comme suit :

TABLEAU II

Nombre de décès par maison	0	1	2	3	4	5
1° Groupe : 28 habitants .	39254	9498	1149	93	6	0
2° Groupe : 48 habitants .	20137	8452	1774	228	26	2

Ce tableau nous montre que la fréquence des cas de cancer, pour un même immeuble augmente beaucoup plus vite que sa population ; il a près de cinq fois plus de maisons à 4 cancers dans les immeubles de 48 habitants que dans les immeubles occupés par 28 personnes seulement. On comprend alors que la fréquence des décès par cancer pour un immeuble donné puisse être beaucoup plus grande quand le nombre des habitants est porté à 80, 100 ou davantage.

Comme nous ignorons la façon dont la population parisienne était répartie dans les 80.639 maisons qu'elle comptait en 1916, il est impossible de préciser, par le calcul des probabilités, la répartition qui nous intéresse, mais, les considérations qui précèdent tendent à nous faire admettre que cette répartition ne doit pas s'écarter notablement des taux réels que la statistique nous a révélés...

CRITIQUE

II. PAR LA METHODE DES STATISTIQUES TEMOINS HETEROGENES. (1)

« Si les arguments que nous venons de développer ne suffisent pas à entraîner la conviction, nous pouvons encore recourir à un autre procédé de démonstration, dont nous avons eu l'idée, en vue d'éliminer plus complètement encore que dans la comparaison précédente, les éléments du problème dont il est difficile d'évaluer les valeurs relatives.

Ce procédé consiste à choisir un phénomène n'ayant aucun rapport avec la cancérose et se répétant en même nombre total dans une même ville, dans les mêmes maisons, avec la même population, et à déterminer les fréquences des répétitions du phénomène par immeuble. C'est la méthode des statistiques témoins, portant sur des phénomènes essentiellement hétérogènes.

Tous les facteurs étant communs, la répartition doit demeurer la même si aucune influence ne vient perturber la règle distributive commandée par le hasard.

Notre statistique des décès par cancer à Lyon ayant porté sur 6.703 cas nous avons pointé, dans les mêmes immeubles, 6.703 naissances consécutives pour mettre ensuite en parallèle les répartitions des deux documents.

La comparaison est enregistrée dans le Tableau III. (2)

(1) AUGUSTE LUMIÈRE. — « La Méthode » des statistiques témoins hétérogènes, CRT 197, 1933 (p. 214).

(2) Nous avons éliminé de notre statistique certaines constructions modernes ou groupes de constructions dépendant d'œuvres sociales où se trouvent réunis des ména-

TABLEAU III

Fréquence par immeuble.	0	1	2	3	4	5	6
Décès par cancer	18231	3769	963	228	53	16	5
Décès totaux	18197	3952	913	210	46	13	4
Naissances	18248	3736	982	212	51	18	6

La concordance est encore ici à tel point remarquable qu'il ne nous est pas possible de représenter graphiquement la ségrégation comparative : les courbes se confondent sensiblement.

Celui qui se laisserait impressionner par la réunion forfuite et inévitable d'un certain nombre de naissances dans les mêmes immeubles, prétendrait-il, en raisonnant comme les partisans des maisons à cancer, qu'il y a des « maisons à naissance », et qu'il y aurait lieu d'en rechercher les causes dans le sol, l'eau, ou les parasites ?

Nous n'insisterons pas sur la puérité d'une telle déduction.

Afin qu'aucun doute ne puisse subsister, en ce qui regarde le rôle du hasard dans la répartition des phénomènes, nous avons encore étudié, de la même manière, la répartition des décès totaux dans la même agglomération et pour le même nombre total de cas (6.703), abstraction faite bien entendu, des hôpitaux ou établissements similaires susceptibles de fausser notre statistique.

ges d'ouvriers en grand nombre et pour lesquels il y a des raisons particulières d'enregistrer des naissances anormalement nombreuses, de même que nous n'avons pas compris dans notre statistique de cancer, les hôpitaux, hospices, maisons de santé où affluent les néoplasiques.

Le Tabelau III, nous montre encore comment ces décès, pour toutes causes et en nombre égal, sont répartis dans les mêmes immeubles.

On voit que la mortalité générale se distribue dans les maisons de la ville, exactement suivant la même loi que les naissances ou la mortalité par cancer, et il en serait certainement de même de tous les autres phénomènes à répétition dont la fréquence par maison ne serait réglée que par le hasard.

On rencontrerait toujours un certain nombre d'immeubles dans lesquels le phénomène se serait produit 4, 5, 6 fois et plus encore, suivant le nombre total de cas considérés. Les coïncidences qui ont été relevées par les cliniciens et qui ont engendré la croyance aux maisons à cancers, ne correspondent donc bien qu'à une très faible fraction de celles qui existent et qui sont commandées par le hasard seul.

Devant les preuves que nous venons de fournir, nous ne pensons pas que l'on puisse désormais parler de maisons à cancers et, en tous cas, si l'on voulait conserver ce vocable, il faudrait invoquer d'autres faits et d'autres arguments que ceux qui sont avancés en sa faveur et que rien n'a pu faire soupçonner jusqu'ici.

La légende des maisons à cancer comporte à notre sens, un enseignement de la plus haute portée ; elle nous montre avec quelle facilité l'observateur accorde sa confiance à des faits que le hasard seul rapproche et l'erreur qu'il commet en prenant de simples coïncidences pour les rapports de cause à effet... »

(Travail du Centre Anticancéreux de Lyon, dirigé par M. le Professeur Bérard.)

En dehors du calcul des probabilités et de la méthode des statistiques témoins hétérogènes que nous venons d'exposer, il existe, pour évaluer le rôle du hasard, une technique expérimentale, dont les résultats facilement contrôlables sont donnés par une méthode fort simple.

Nous empruntons à Auguste Lumière, dans le *Bulletin de l'Académie de Médecine*, Tome CXV, 1935 (p. 23), l'exposé de cette méthode appliquée au cancer familial.

CRITIQUE

III. PAR LA METHODE EXPERIMENTALE

« Mélangeons dans un sac ou dans un récipient quelconque, 1.000 billes noires et 14.000 billes blanches, ces nombres représentant respectivement la proportion des décès par cancer et des décès pour tout autre cause ; procédons ensuite à un grand nombre de tirages, effectués au hasard, de groupes de 10 billes par exemple, représentant des familles de 10 membres et notons pour chaque tirage le nombre de billes de chaque sorte.

Dans une première expérience, ainsi réalisée et portant sur 4.000 prélèvements de 10 billes, nous avons trouvé 1935 prélèvements où toutes les billes étaient blanches ; 479 avec 2 noires et 8 blanches ; 107 avec 3 noires et 7 blanches ; 17 avec 4 noires et 6 blanches et 4 avec 5 noires et 5 blanches.

Par conséquent, du seul fait du hasard, si nous prenons 4.000 familles de 10 membres, nous devons en trouver 17 où quatre de ces membres sur dix doivent mourir de cancer, et quatre dans lesquels la moitié de

la famille doit succomber de la même affection.

Ces dénombrements n'ont rien d'absolu ; pour qu'ils aient une valeur intégrale, il faudrait que leur nombre soit beaucoup plus considérable, mais cependant on peut affirmer que déjà, avec 4.000 tirages, nous nous rapprochons beaucoup des proportions réelles.

Nous avons répété cette expérience sur des groupes de 5 et 7 personnes et le résultat de nos 4.000 tirages dans ces différents cas, est indiqué dans le tableau ci-après.

Nombre de billes par tirage	Nombre de billes noires					
	0	1	2	3	4	5
5.....	2829	1010	151	10	0	0
7.....	2273	1319	354	47	6	1
10.....	1935	1458	479	107	17	4

Sans que l'hérédité joue le moindre rôle dans la prorogation de la cancérose, par le seul jeu des lois du hasard on devra donc rencontrer des familles dans lesquelles plusieurs membres seront fatalement victimes de l'affection.

Il n'est pas possible qu'il en soit autrement et nous avons déterminé les proportions approximatives dans lesquelles ces répétitions devaient se produire.

On ne peut donc pas affirmer que les exemples cités dans la littérature médicale correspondent à autre chose qu'à de simples coïncidences. Pour fixer la question, il est indispensable de dresser des statistiques bien contrôlées, portant sur un grand nombre de familles de même composition numérique et examiner si les fréquences des répétitions dépassent celles du tableau ci-dessus.

Les états démographiques publiés jusqu'ici, sans tenir compte de ces remarques n'ont absolument aucune signification et les conclusions qu'on en a tirées sont entachées d'erreur... »

CONCLUSION

En lisant ce travail et celui très voisin du Docteur Charlette Ricard, le Médecin, praticien, voulant arriver à des conclusions sur la valeur d'un thérapeutique, ou épidémiologiste désirant apprécier le résultat de telle ou telle méthode prophylactique, sera peut-être déçu par les exigences des statisticiens.

Des travaux si méticuleux, des conditions si nombreuses qui semblent ne pas permettre, de conclusions fermes, l'inciteraient presque à abandonner la méthode des statistiques.

Une telle déception nous paraît aussi peu justifiée que la satisfaction de ceux qui croient fermement à la signification d'un écart de 2 % dans un pourcentage établi à partir de 50 éléments.

Nous avons voulu proposer, à ceux qui se servent des statistiques, une méthode leur permettant de ne pas tomber dans l'un de ces deux excès.

Cette méthode pourra se résumer sous forme de deux conseils qui seront les conclusions de notre travail :

- 1° Dans tout travail statistique, tendant à établir une proportion quelconque d'éléments, il faut toujours tenir compte du fait qu'il existe un écart, entre le chiffre absolu recherché et celui donné par la méthode employée. Cet écart, (écart probable ou marge d'erreur inévitable), sera d'autant

plus grand que le nombre d'éléments retenus pour les calculs sera plus petit, par rapport à la masse totale des éléments en étude. Seules des différences supérieures à la somme des écarts probables permettent d'affirmer qu'il y a eu modification.

- 2° Lorsqu'une statistique tend à montrer une distribution inégale de faits donnés, on n'est en droit d'y voir un rapport quelconque avec l'un des facteurs, que lorsque le calcul des probabilités montre que cette distribution n'est pas voisine de celle du hasard.
-

BIBLIOGRAPHIE

- BERTRAND (J.). — *Calcul des probabilités*, édité chez Gauthier Vilars, in 8 (25-16) 2^e édition 1917.
- BESSON. — *Le cancer et les maisons*, Annales d'hygiène publique, industrielle et sociale. Tome I. 1923 (p. 23).
- BRADFORD HILL D. Sc. PH. D. — *Principles of medical statistics*, published by « The Lancet » limited 7 Adam street adelphi London. 1937.
- CHATON. — *De l'intérêt très particulier de l'étude des maisons à cancer et de la manière dont on pourrait y procéder*, Bulletin de l'Académie de Médecine. Tome 107, 1932 (p. 875).
- CARVALLO (E.). — *Le calcul des probabilités et ses applications*, in 8 (25-16) Gauthiers Villars, édit. 1912.
- FAVRE. — *De la Contagion du cancer*, Thèse Lyon 1892.
- HERBER (J.). — *Le paludisme au camp de.... (Maroc), ses causes*. Presse médicale du 19 Mai 1920.
- LAPLACE. — *Théorie analytique des probabilités*. 1812.
- LAPLACE. — *Œuvres complètes*. Tome X (p. 316 et suivantes). Tome VII, 2^e partie (p. 388). Tome XIV (p. 165).
- LUMIERE (A.). — *La contagion, l'hérédité et le hasard dans le cancer*. Bulletin de l'Académie de Médecine, Tome 98, 1927 (p. 89).
- LUMIERE (A.). — *La méthodes des statistiques témoins hé-térogènes*. CRT 197, 1933 (p. 214).
- LUMIERE (A.) et PAUL VIGNES. — *Au sujet des maisons à cancer*. Avenir Médical, XXX^e année, Février 1933.
- LUMIERE (A.). — *Bulletin de l'Académie de Médecine*, Tome CXV, 1935 (p. 23).

- MATHIEU (J.). — *L'infection tuberculeuse révélée par les cuti-réactions à la tuberculine, chez les indigènes de Figuig (Maroc Oriental)*. Archives des Instituts Pasteur de l'Afrique du Nord, 1928, Tome VI (p. 18 à 22).
- MILLOT (S.). — *Théorie nouvelle de la probabilité des causes*. Gauthier Villars édit. 1925.
- MONTESUS DE BALLORE (R.). — *Leçons élémentaires sur le calcul des probabilités*. Gauthier Villars édit. 1908.
- OCAGNE (M. d'). — Notes présentées à l'Académie des Sciences. 26 - XII - 22 ; 23 - IV - 23 ; 11 - VI - 23. Insérées aux comptes rendus. Tome 176, 1923 (p. 30, 564, 1126, 1695).
- POINCARÉ (H.). — *Calcul des probabilités*, Gauthier Villars édit. 1923.
- ROSS AND SOTT. — *Tables of statistical error*. The Annals of tropical medicine and parasitology. T. 5, 24 October 1911 (p. 347).
- SABRIE. — Archives des Instituts Pasteur de l'Afrique du Nord. Tome II, 1922 (p. 336).
- SENEVET (G.) et WITAS (P.). — *Action sur certains sérums d'un mélange de sublimé et de sulfate de soude en solution aqueuse*. Archives des Instituts Pasteur de l'Afrique du Nord. Tome XI, 1922 (p. 457).

Vu :

Le Président de thèse

SENEVET

Vu :

Le Doyen :

LEBLANC

Vu et permis d'imprimer,

Alger, le 25 Mars 1938

POUR LE RECTEUR ET PAR ORDRE,

Le Doyen de la Faculté des Sciences :

ROUYER.

Tables des Matières

	PAGES
Avant propos	17
Observation N° 1	19
<i>(Exemple d'erreur par la méconnaissance de la notion d'écart probable.)</i>	
Observation N° 2	24
<i>(Exemple d'erreur par la méconnaissance du rôle d'hasard dans les distributions inégales des faits.)</i>	
CHAPITRE I	
Qu'est-ce que le rôle du hasard ?.....	25
CHAPITRE II	
Les principes du calcul des probabilités	33
CHAPITRE III	
La notion d'écart probable ou marge d'erreur inévitable	45
(A) Le calcul de l'écart probable par l'application des	
Tables de Ross et Scott	49
Table A 1 (99.999 contre 1)	76
Table A 2 (9.999 contre 1)	77
Table A 3 (999 contre 1)	78
Table A 4 (99 contre 1)	79
Table A 5 (9 contre 1)	80
Table A 6 (1 contre 1)	81
(B) Le calcul de l'écart probable par l'emploi des nomogrammes	83
Nomogrammes pour l'application du théorème de	
Jacques Bernouilli	87
Premier exemple de calcul simplifié	88
Deuxième exemple de calcul simplifié	91
CHAPITRE IV	
La notion d'écart probable montre l'inanité des décimales dans les pourcentages grands que l'unité	97
I. Les décimales dans les pourcentages plus grands que un	100
II. Les décimales dans les pourcentages plus petits que un	101
CHAPITRE V	
Le rôle du hasard dans l'interprétation des concomitances	103
Critique par le calcul des probabilités	107
Critique par la méthode des statistiques témoins hétérogènes	110
Critique par la méthode expérimentale	113
Conclusion	117
Bibliographie	119

